Some numerical aspects of Arnoldi-Tikhonov regularization

Maged Alkilayh[a], Lothar Reichel[b]

[a] *Department of Mathematical Sciences, Kent State University, Kent, OH 44242, USA, and Department of Mathematics, College of Science, Qassim University, Buraydah 51452, Saudi Arabia.*

[b] *Department of Mathematical Sciences, Kent State University, Kent, OH 44242, USA.*

Abstract

Large linear discrete ill-posed problems are commonly solved by first reducing them to small size by application of a few steps of a Krylov subspace method, and then applying Tikhonov regularization to the reduced problem. A regularization parameter determines how much the given problem is regularized before solution. If the matrix of the linear discrete ill-posed problem is nonsymmetric, then the reduction often is carried out with the aid of the Arnoldi process, while when the matrix is symmetric, the Lanczos process is used. This paper discusses several numerical aspects of these solution methods. We illustrate that it may be beneficial to apply the Arnoldi process also when the matrix is symmetric and discuss how certain user-chosen basis vectors can be added to the solution subspace. Finally, we compare the application of the discrepancy principle for the determination of the regularization parameter to another approach that is based on the solution of a cubic equation and has been proposed in the literature.

Keywords: ill-posed problem, regularization, Arnoldi decomposition, discrepancy principle.
2010 MSC: 65F10, 65R30

1. Introduction

We consider the problem of computing an approximate solution of linear systems of equations

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

with a large square symmetric or nonsymmetric matrix $A \in \mathbb{R}^{n\times n}$, whose singular values "cluster" at the origin, i.e., A has many singular values of different orders of magnitude close to the origin. In particular, A is severely ill-conditioned

Email addresses: malkilay@kent.edu (Maged Alkilayh), reichel@math.kent.edu (Lothar Reichel)

and may be rank-deficient. Linear systems of equations with a matrix of this kind are commonly referred to as linear discrete ill-posed problems. They arise, for instance, from the discretization of linear ill-posed problems, such as Fredholm integral equations of the first kind with a square integrable kernel. The right-hand side $\boldsymbol{b} \in \mathbb{R}^n$ of many linear discrete ill-posed problems (1) that arise in applications, e.g., in engineering and physics, represents available data, and typically is contaminated by an unknown error $\boldsymbol{e} \in \mathbb{R}^n$. This error may stem from measurement inaccuracies and discretization.

Let $\hat{\boldsymbol{b}} \in \mathbb{R}^n$ denote 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 unavailable linear system of equations

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

to be consistent, i.e., $\hat{\boldsymbol{b}}$ is in the range of A. This condition should be satisfied when applying the discrepancy principle for determining the regularization parameter in Tikhonov regularization; see below. Other methods for determining this parameter also will be commented on.

Let $\hat{\boldsymbol{x}} \in \mathbb{R}^n$ denote the solution of minimal Euclidean norm of (3). We would like to determine an accurate approximation of $\hat{\boldsymbol{x}}$ by computing a suitable approximate solution of the available linear system (1) with the error-contaminated right-hand side $\boldsymbol{b}$ with the aid of an iterative method.

Straightforward solution of (1) generally does not yield a meaningful approximation of $\hat{\boldsymbol{x}}$ due to the ill-conditioning of A and the error $\boldsymbol{e}$ in $\boldsymbol{b}$. Therefore, one often replaces (1) by a minimization problem whose solution is less sensitive to the error $\boldsymbol{e}$ than the solution of (1), and solves the minimization problem instead of (1). This replacement is known as regularization.

One of the most popular regularization methods is due to Tikhonov. This method replaces (1) by the penalized least-squares problem

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

where $L \in \mathbb{R}^{s\times n}$ is a regularization matrix and $\mu > 0$ is a regularization parameter. The matrix L often is chosen to be the identity matrix I, or a discretized differential operator such that the null spaces of A and L only intersect at the origin. For ease of discussion, we will assume that $L = I$, but we comment below on how more general regularization matrices L can be handled. 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 $\hat{\boldsymbol{x}}$. Throughout this paper $\|\cdot\|$ denotes the Euclidean vector norm.

This paper discusses several numerical aspects of solving large-scale Tikhonov regularization problems (4). When the matrix $A \in \mathbb{R}^{n\times n}$ is nonsymmetric, the Arnoldi process is a popular approach to reduce the matrix A in (4) to a matrix of small size; see, e.g., [6, 11, 12, 22]. This is described in Section 2. The use of

the Arnoldi process is particularly attractive when matrix-vector products with the matrix A are fairly inexpensive to evaluate, but matrix-vector products with the transpose of A, denoted by A^T, are not. This situation may arise when the matrix A is not explicitly stored such as when solving large nonlinear systems of equations by Krylov subspace methods; see, e.g., [8] for a discussion. It also may be difficult to evaluate matrix-vector products with A^T when matrix-vector products with A are evaluated by a multipole method; see, e.g., [13]. Moreover, iterative methods that are based on the Arnoldi process may require fewer matrix-vector product evaluations than iterative methods that demand matrix-vector product evaluations with both A and A^T. An illustration of the latter is provided in [5], where the restoration of large images is described.

When the matrix $A \in \mathbb{R}^{n\times n}$ is symmetric, the Arnoldi process can be replaced by the symmetric Lanczos process, which also is discussed in Section 2. Application of ℓ steps of the Lanczos process requires less arithmetic work than application of the same number of steps of the Arnoldi process. However, we illustrate in Section 5 that reduction of A by ℓ steps of the Arnoldi process may result in computed approximations of the solution of (4) of significantly higher quality than when A is reduced by ℓ steps of the Lanczos process. This leads to the observation that when solving large-scale linear discrete ill-posed problems with a symmetric matrix, the matrix should be reduced to a smaller matrix by the Arnoldi process.

The regularization parameter $\mu > 0$ and the number of steps ℓ of the Arnoldi process can be determined with the aid of the discrepancy principle, which is reviewed in Section 2. Specifically, the number of steps is chosen large enough so that the discrepancy principle can be satisfied. We find that the quality of the computed solution may be improved by letting the number of steps be slightly larger than the smallest number of steps needed to satisfy the discrepancy principle. Increasing the number of steps further typically does not result in additional improvement of the quality of the computed solution. Details are provided in Section 2. In Section 3 we describe a method for determining the regularization parameter proposed by Neubauer [20], who considers the situation when both the matrix A and the vector $\boldsymbol{b}$ are contaminated by error; we are concerned with the special case when only $\boldsymbol{b}$ is error-contaminated. We find the discrepancy principle to determine a regularization parameter-value that gives an approximation of $\hat{\boldsymbol{x}}$ of higher quality than when applying Neubauer's method.

For some linear discrete ill-posed problems (1), the solution subspace obtained by carrying out ℓ steps of the Arnoldi process with ℓ determined as described above does not furnish a solution subspace that is well suited for approximating the desired solution $\hat{\boldsymbol{x}}$ accurately. In this situation, it may be beneficial to augment the solution subspace determined by the Arnoldi process by a user-chosen subspace that allows the representation of important features of $\hat{\boldsymbol{x}}$. The flexible Arnoldi process introduced by Saad [25] provides an approach for choosing a fairly arbitrary solution subspace. This process is used in the flexible GMRES method, whose application to the solution of large-scale linear discrete ill-posed problems is described in [19]. A drawback of the approach discussed

in [19] is that the determination of a solution subspace of dimension ℓ requires storage of about 2ℓ vectors in $\mathbb{R}^n$, while the generation of an ℓ-dimensional solution subspace with the standard Arnoldi process only demands storage of about ℓ vectors in $\mathbb{R}^n$. Section 4 describes how a few user-chosen vectors can be included in a standard Krylov subspace with modest storage requirement. A few computed examples are presented in Section 5, and concluding remarks with comments on an extension can be found in Section 6.

The computed examples illustrate applications of the discrepancy principle and Neubauer's method for determining the regularization parameter. These methods require a bound for $\|\boldsymbol{e}\|$ to be explicitly known. Findings of this paper also apply when other techniques for determining a suitable value of the regularization parameter are used that do not require a bound for $\|\boldsymbol{e}\|$ to be available. These techniques include the L-curve criterion, generalized cross validation, and cross validation; see [4, 12, 15, 17, 18, 23] for discussions of such methods.

2. The discrepancy principle and Krylov subspace methods

Assume that a fairly accurate bound δ for the norm of the error $\boldsymbol{e}$ in $\boldsymbol{b}$ is known, i.e.,

$$\|\boldsymbol{e}\| \leq \delta, \tag{5}$$

and note that for any $\mu > 0$ the solution of (4) is given by

$$\boldsymbol{x}_\mu = (A^T A + \frac{1}{\mu} I)^{-1} A^T \boldsymbol{b}. \tag{6}$$

The discrepancy principle prescribes that $\mu > 0$ be determined so that

$$\|A\boldsymbol{x}_\mu - \boldsymbol{b}\| = \eta\delta, \tag{7}$$

where $\eta \geq 1$ is a user-chosen constant that is independent of δ. It is shown by Engl et al. [10] in a Hilbert space setting that when (3) is consistent, $\eta > 1$, and μ is determined by (7), the Tikhonov solution $\boldsymbol{x}_\mu$ converges to $\hat{\boldsymbol{x}}$ as $\delta \searrow 0$.

The solution of (7) and the computation of the Tikhonov solution (6) is expensive or unfeasible when A is a large matrix. We therefore reduce A to a matrix $H_{\ell+1,\ell} \in \mathbb{R}^{(\ell+1)\times\ell}$ of small size by applying ℓ steps of the Arnoldi process to A with initial vector $\boldsymbol{v}_1 = \boldsymbol{b}/\|\boldsymbol{b}\|$. This is summarized in Algorithm 1. Generically, this yields the Arnoldi decomposition

$$AV_\ell = V_{\ell+1} H_{\ell+1,\ell}, \tag{8}$$

where the matrix $V_{\ell+1} = [\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_{\ell+1}] \in \mathbb{R}^{n\times(\ell+1)}$ has orthonormal columns, V_ℓ consisting of the first ℓ columns of $V_{\ell+1}$, and $H_{\ell+1,\ell}$ is of upper Hessenberg form; i.e., all entries below the subdiagonal vanish. The range of V_ℓ is the Krylov subspace

$$\mathbb{K}_\ell(A, \boldsymbol{b}) = \text{span}\{\boldsymbol{b}, A\boldsymbol{b}, \ldots, A^{\ell-1}\boldsymbol{b}\}, \tag{9}$$

which will be the solution subspace; see, e.g., Saad [26] for further details on the Arnoldi process.

Algorithm 1 (The Arnoldi Process).
Input: *matrix A, vector $\boldsymbol{b} \neq \boldsymbol{0}$, number of steps ℓ;*

1. *Initialize:* $\boldsymbol{v}_1 = \boldsymbol{b}/\|\boldsymbol{b}\|$
2. ***for*** $j = 1, \ldots, \ell$ ***do***
3. $\quad \boldsymbol{w} = A\boldsymbol{v}_j$
4. $\quad$ ***for*** $i = 1, \ldots, j$ ***do***
5. $\quad\quad h_{i,j} = \boldsymbol{v}_i^T \boldsymbol{w}$;
6. $\quad\quad \boldsymbol{w} = \boldsymbol{w} - h_{i,j}\boldsymbol{v}_i$;
7. $\quad$ ***end for***
8. $\quad h_{j+1,j} = \|\boldsymbol{w}\|$;
9. $\quad$ ***if*** $h_{j+1,j} = 0$ ***then*** *stop*
10. $\quad \boldsymbol{v}_{j+1} = \boldsymbol{w}/h_{j+1,j}$;
11. ***end for***

Output: *The matrices V_ℓ, $V_{\ell+1}$, and $H_{\ell+1,\ell}$ in (8).*

The Arnoldi process breaks down at step j if $h_{j+1,j} = 0$. Then the computations simplify. Since this event is very rare, we will not dwell on the details.

The Arnoldi-Tikhonov (AT) method replaces the minimization problem (4) by

$$\min_{\boldsymbol{x} \in \mathbb{K}_\ell(A,\boldsymbol{b})} \{\|A\boldsymbol{x} - \boldsymbol{b}\|^2 + \frac{1}{\mu}\|\boldsymbol{x}\|^2\} \tag{10}$$

for a suitable value of ℓ. Substituting

$$\boldsymbol{x}_\ell = V_\ell \boldsymbol{y}_\ell$$

into (10) and using the decomposition (8) yields the reduced minimization problem

$$\min_{\boldsymbol{y} \in \mathbb{R}^\ell} \{\|H_{\ell+1,\ell}\boldsymbol{y} - \boldsymbol{e}_1\|\boldsymbol{b}\|\|^2 + \frac{1}{\mu}\|\boldsymbol{y}\|^2\}, \tag{11}$$

which has a unique solution $\boldsymbol{y}_{\mu,\ell}$ for any $\mu > 0$. Then

$$\boldsymbol{x}_{\mu,\ell} = V_\ell \boldsymbol{y}_{\mu,\ell}$$

solves (10), and furnishes an approximate solution of (4).

The discrepancy principle prescribes that $0 < \mu < \infty$ be determined so that $\boldsymbol{y}_{\mu,\ell}$ satisfies

$$\|H_{\ell+1,\ell}\boldsymbol{y}_{\mu,\ell} - \boldsymbol{e}_1\|\boldsymbol{b}\|\| = \eta\delta. \tag{12}$$

This requires that

$$\min_{\boldsymbol{y} \in \mathbb{R}^\ell} \|H_{\ell+1,\ell}\boldsymbol{y} - \boldsymbol{e}_1\|\boldsymbol{b}\|\| < \eta\delta. \tag{13}$$

We will comment on this inequality below.

Proposition 2.1. *Assume that $\ell+1$ steps of the Arnoldi process applied to A with initial vector $\boldsymbol{b}$ can be carried out. Then*

$$\min_{\boldsymbol{y}\in\mathbb{R}^{\ell+1}} \|H_{\ell+2,\ell+1}\boldsymbol{y} - \boldsymbol{e}_1\|\boldsymbol{b}\|\| \leq \min_{\boldsymbol{y}\in\mathbb{R}^{\ell}} \|H_{\ell+1,\ell}\boldsymbol{y} - \boldsymbol{e}_1\|\boldsymbol{b}\|\|. \tag{14}$$

Assume that $H_{\ell+1,\ell}^T\boldsymbol{e}_1 \neq \boldsymbol{0}$. Then the function

$$\phi_\ell(\mu) = \|H_{\ell+1,\ell}\boldsymbol{y}_{\mu,\ell} - \boldsymbol{e}_1\|\boldsymbol{b}\|\|^2 \tag{15}$$

is strictly decreasing and strictly convex for $\mu \geq 0$ with $\phi_\ell(0) = \|\boldsymbol{b}\|^2$.

Proof: The inequality (14) follows from fact that the minimization problem on the right-hand side is obtained by setting the last entry of the vector $\boldsymbol{y}$ in the minimization problem on the left-hand side to zero. Typically, the inequality is strict.

The stated properties of the function (15) can be shown by using the fact that the solution $\boldsymbol{y}_{\mu,\ell}$ of (11) can be written as

$$\boldsymbol{y}_{\mu,\ell} = (H_{\ell+1,\ell}^T H_{\ell+1,\ell} + \mu^{-1} I)^{-1} H_{\ell+1,\ell}^T \boldsymbol{e}_1 \|\boldsymbol{b}\|. \tag{16}$$

Substituting this expression into (15) gives

$$\phi_\ell(\mu) = \|\boldsymbol{b}\|^2 \|(H_{\ell+1,\ell}(H_{\ell+1,\ell}^T H_{\ell+1,\ell} + \mu^{-1} I)^{-1} H_{\ell+1,\ell}^T - I)\boldsymbol{e}_1\|^2. \tag{17}$$

Consider the singular value decomposition (SVD)

$$H_{\ell+1,\ell} = U_{\ell+1}\Sigma_{\ell+1,\ell}W_\ell^T, \tag{18}$$

where the matrices $U_{\ell+1} \in \mathbb{R}^{(\ell+1)\times(\ell+1)}$ and $W_\ell \in \mathbb{R}^{\ell\times\ell}$ are orthogonal, and the matrix

$$\Sigma_{\ell+1,\ell} = \mathrm{diag}[\sigma_1, \sigma_2, \ldots, \sigma_\ell] \in \mathbb{R}^{(\ell+1)\times\ell}$$

is diagonal with nonnegative diagonal entries σ_j (the singular values) in decreasing order. Substituting the SVD into (17) gives

$$\phi_\ell(\mu) = \|\boldsymbol{b}\|^2 \boldsymbol{e}_1^T U_{\ell+1} \mathrm{diag}[(\mu\sigma_1^2+1)^{-2}, (\mu\sigma_2^2+1)^{-2}, \ldots, (\mu\sigma_\ell^2+1)^{-2}, 1] U_{\ell+1}^T \boldsymbol{e}_1, \tag{19}$$

which shows that $\phi_\ell'(\mu) \leq 0$ and $\phi_\ell''(\mu) \geq 0$. Equality can be achieved only if the vector (16) is independent of μ. This can only happen if $H_{\ell+1,\ell}^T\boldsymbol{e}_1 = \boldsymbol{0}$. It follows that $\phi(\mu)$ is strictly decreasing and strictly convex for $\mu > 0$. The expression (19) shows that $\phi_\ell(\mu)$ can be continuously extended to $\mu = 0$. □

Proposition 2.2. *Assume the conditions of Proposition 2.1 hold, and let all subdiagonal entries $h_{j+1,j}$, $1 \leq j \leq \ell$, of the upper Hessenberg matrix $H_{\ell+1,\ell}$ be positive. Then*

$$\lim_{\mu\to\infty} \phi_\ell(\mu) = \|\boldsymbol{b}\|^2 \|\boldsymbol{e}_{\ell+1}^T U_{\ell+1}^T \boldsymbol{e}_1\|^2, \tag{20}$$

where $U_{\ell+1}$ is the left orthogonal matrix in the SVD of $H_{\ell+1,\ell}$; see (18).

Proof: When applying ℓ steps of the Arnoldi process to the matrix A with initial vector $\boldsymbol{b}$, generically, all subdiagonal entries of the upper Hessenberg matrix $H_{\ell+1,\ell} \in \mathbb{R}^{(\ell+1)\times\ell}$ in the Arnoldi decomposition (8) are positive. We will comment on the rare nongeneric situation below. Therefore, generically, the columns of the matrix $H_{\ell+1,\ell}$ are linearly independent and then all the singular values $\sigma_1, \sigma_2, \ldots, \sigma_\ell$ of $H_{\ell+1,\ell}$ are positive. The limit (20) is a consequence of (19).

When applying ℓ steps of the Arnoldi process to the matrix A with initial vector $\boldsymbol{b}$ without breakdown, the first ℓ subdiagonal entries of the upper Hessenberg matrix $H_{\ell+1,\ell}$ are positive, but the last subdiagonal entry, $h_{\ell+1,\ell}$ may vanish. Then the first $\ell-1$ columns of $H_{\ell+1,\ell}$ are linearly independent, and it follows that the first $\ell-1$ singular values of $H_{\ell+1,\ell}$ are nonvanishing. We then obtain the bound

$$\lim_{\mu\to\infty} \phi_\ell(\mu) \leq \|\boldsymbol{b}\|^2 \|(\boldsymbol{e}_{\ell+1}^T + \boldsymbol{e}_\ell^T) U_{\ell+1}^T \boldsymbol{e}_1\|^2.$$

□

We turn to the inequality (13). Assume that (12) holds for a finite $\mu_\ell = \mu > 0$. Since the function (15) is strictly decreasing, we have $\phi_\ell(2\mu_\ell) < \phi_\ell(\mu_\ell)$. Hence, there is a vector $\boldsymbol{y} = \boldsymbol{y}_{2\mu_\ell,\ell}$ such that $\phi_\ell(2\mu_\ell) < \phi_\ell(\mu_\ell)$. Consequently, inequality (13) holds.

For ease of discussion, we assume the generic case to hold in Proposition 2.2. Equation (12) can be expressed as

$$\phi_\ell(\mu) = \eta^2\delta^2, \tag{21}$$

and we assume that

$$\eta^2\delta^2 > \|\boldsymbol{b}\|^2 \|\boldsymbol{e}_{\ell+1}^T U_{\ell+1}^T \boldsymbol{e}_1\|^2.$$

Then there is a finite smallest index ℓ such that equation (21) can be satisfied. We denote this index by $\ell_{\rm dis}$. For any $\ell \geq \ell_{\rm dis}$, we denote the solution of (21) by μ_ℓ. When $\eta^2\delta^2/\|\boldsymbol{b}\|^2$ is not very small, the value of $\ell_{\rm dis}$ typically is not large. This is illustrated in Section 5. The computed examples of Section 5, as well as many other computed examples, indicate that it often is beneficial to carry out slightly more than $\ell_{\rm dis}$ steps of Algorithm 1. We will see that carrying out $\ell = \ell_{\rm dis} + 2$ steps often is beneficial.

An efficient way to determine $\ell_{\rm dis}$ is to solve

$$\min_{\boldsymbol{y}\in\mathbb{R}^\ell} \|H_{\ell+1,\ell}\boldsymbol{y} - \boldsymbol{e}_1\|\boldsymbol{b}\|\|, \tag{22}$$

say by QR factorization of $H_{\ell+1,\ell}$ for increasing values of ℓ, by updating the available QR factorization of $H_{\ell,\ell-1}$ (see, e.g., [9] for details), and selecting $\ell_{\rm dis}$ as the smallest value of ℓ such that the solution of (22) satisfies (13).

We determine μ_ℓ with a zero-finder, such as Newton's method. This method requires that values of ϕ_ℓ and its first derivative $\phi_\ell'(\mu)$ with respect to μ be computed at approximations $\mu_\ell^{(j)}$ of μ_ℓ for $j = 0, 1, 2, \ldots$. Having computed the SVD of $H_{\ell+1,\ell}$, see (18), the expression (19) can be evaluated in only $O(\ell)$

arithmetic floating point operations (flops) for each value of μ. The derivative $\phi'_\ell(\mu)$ has a representation analogous to (19).

Having computed μ_ℓ, we determine the associated vector (16) by solving the least-squares problem

$$\min_{\boldsymbol{y}\in\mathbb{R}^{\ell+1}} \left\| \begin{bmatrix} H_{\ell+1,\ell} \\ \mu_\ell^{-1/2} I \end{bmatrix} \boldsymbol{y} - \begin{bmatrix} \boldsymbol{e}_1\|\boldsymbol{b}\| \\ \boldsymbol{0} \end{bmatrix} \right\|, \tag{23}$$

because this avoids solving the system of equations

$$(H_{\ell+1,\ell}^T H_{\ell+1,\ell} + \mu_\ell^{-1} I)\boldsymbol{y} = \boldsymbol{e}_1\|\boldsymbol{b}\|. \tag{24}$$

This can be important when $1/\mu_\ell$ is tiny, because the matrix $H_{\ell+1,\ell}$ may be ill-conditioned. If the SVD of $H_{\ell+1,\ell}$ is available, then it can be substituted into (23).

Our reason for using the factor $1/\mu$ instead of μ in (4) and (11) is that this makes $\phi_\ell(\mu)$ decreasing and convex for $\mu > 0$. This implies that Newton's method converges to the solution μ_ℓ monotonically and quadratically when the initial approximation $\mu_\ell^{(0)}$ of μ_ℓ is chosen smaller than or equal to μ_ℓ. In the computed examples of Section 5, we let $\mu_\ell^{(0)} = 0$.

We remark that other zero-finders than Newton's method also can be used, see, e.g., [3, 24], but they have to be safe-guarded in order to secure convergence. Their application therefore is more complicated. This also holds if Newton's method is applied to determine the regularization parameter for (11) with μ replaced by $1/\mu$.

When the matrix A is symmetric, the decomposition (8) simplifies to the Lanczos decomposition

$$AV_\ell = V_{\ell+1}T_{\ell+1,\ell}, \tag{25}$$

where $V_{\ell+1} = [\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_{\ell+1}] \in \mathbb{R}^{n\times(\ell+1)}$ has orthonormal columns with $\boldsymbol{v}_1 = \boldsymbol{b}/\|\boldsymbol{b}\|$, the matrix $V_\ell \in \mathbb{R}^{n\times\ell}$ consists of the first ℓ columns of $V_{\ell+1}$, and

$$T_{\ell+1,\ell} = \begin{bmatrix} \alpha_1 & \beta_2 & & & & & \\ \beta_2 & \alpha_2 & \beta_3 & & & & \\ & \beta_3 & \alpha_3 & \ddots & & & \\ & & \ddots & \ddots & \beta_{\ell-1} & & \\ & & & \beta_{\ell-1} & \alpha_{\ell-1} & \beta_\ell & \\ & & & & \beta_\ell & \alpha_\ell & \\ & & & & & \beta_{\ell+1} & \end{bmatrix} \in \mathbb{R}^{(\ell+1)\times\ell}$$

is a tridiagonal matrix with a leading symmetric $\ell \times \ell$ submatrix. The Arnoldi process simplifies to the Lanczos process, which is described by Algorithm 2. Thus, ℓ steps of the Lanczos process determine the diagonal and subdiagonal elements of $T_{\ell+1,\ell}$. Generically, the subdiagonal entries β_j are positive, and then the decomposition (25) with the stated properties exists.

Algorithm 2 (The Symmetric Lanczos Process).
Input: *matrix A, vector $\boldsymbol{b} \neq \boldsymbol{0}$, number of steps ℓ;*

1. *Initialize:* $\boldsymbol{v}_1 = \boldsymbol{b}/\|\boldsymbol{b}\|$, $\beta_1 = 0$, $\boldsymbol{v}_0 = 0$;
2. ***for*** $j = 1, \ldots, \ell$ ***do***
3. $\quad \boldsymbol{w} = A\boldsymbol{v}_j - \beta_j \boldsymbol{v}_{j-1}$;
4. $\quad \alpha_j = \boldsymbol{v}_j^T \boldsymbol{w}$;
5. $\quad \boldsymbol{w} = \boldsymbol{w} - \alpha_j \boldsymbol{v}_j$;
6. $\quad \beta_{j+1} = \|\boldsymbol{w}\|$;
7. $\quad$ ***if*** $\beta_{j+1} = 0$ ***then*** *stop*
8. $\quad \boldsymbol{v}_{j+1} = \boldsymbol{w}/\beta_{j+1}$;
9. ***end***

Output: *The matrices V_ℓ, $V_{\ell+1}$, and $T_{\ell+1,\ell}$ in (25).*

Generally, the Lanczos process is used instead of the Arnoldi process to reduce a large symmetric matrix A to a small matrix, because of its lower computational complexity: ℓ steps of the Lanczos process require about $4\ell n$ flops, while ℓ steps of the Arnoldi process demand about $2\ell^2 n$ flops, in addition to the arithmetic work necessary to evaluate ℓ matrix-vector products with A. However, we will see in Section 5 that reduction of a symmetric matrix A to a small matrix by the Arnoldi process gives approximate solutions of higher quality than when A is reduced by the Lanczos process. This depends on loss of orthogonality of the vectors $\boldsymbol{v}_j$ determined by the Lanczos process in finite-precision arithmetic; the Arnoldi process is equivalent to the Lanczos process with reorthogonalization. Algorithm 3 summarizes the computations of the Arnoldi-Tikhonov method.

Algorithm 3 (The Arnoldi-Tikhonov (AT) algorithm).
Input: *matrix A, vector $\boldsymbol{b} \neq \boldsymbol{0}$, parameters $\eta \geq 1$, $\delta > 0$, initial number of steps ℓ_{init}, concluding number of steps ℓ_{end};*

1. *Let $\ell = \ell_{\text{init}}$;*
2. *Compute $V_{\ell+1}$ and $H_{\ell+1,\ell}$ by Algorithm 1;*
3. *Solve the minimization problem*

$$\min_{\boldsymbol{y} \in \mathbb{R}^\ell} \|H_{\ell+1,\ell}\boldsymbol{y} - \boldsymbol{e}_1\|\boldsymbol{b}\|\|;$$

4. ***if*** $\|H_{\ell+1,\ell}\boldsymbol{y}_\ell - \boldsymbol{e}_1\|\boldsymbol{b}\|\| \geq \eta\delta$; ***then***
5. $\quad \ell = \ell + 1$; ***goto*** *step 2;*
6. ***end if***
7. $\ell_{dis} = \ell$;
8. *Carry out ℓ_{end} additional steps with Algorithm 1, using the available matrices $V_{\ell+1}$ and $H_{\ell+1,\ell}$ as input.*
9. *Let $i = \ell + \ell_{\text{end}}$. We have computed the matrices V_{i+1} and $H_{i+1,i}$;*
10. *Determine the regularization parameter μ_i by solving equation (21) with ℓ replaced by i;*

11. *Solve the minimization problem*

$$\min_{\boldsymbol{y}\in\mathbb{R}^i}\{\|H_{i+1,i}\boldsymbol{y}-\boldsymbol{e}_1\|\boldsymbol{b}\|\|^2+\frac{1}{\mu_i}\|\boldsymbol{y}\|^2\};$$

Output*: Approximate solution* $\boldsymbol{x}_{\mu_i,i}$*,* ℓ_{dis}*,* i*.*

We conclude this section with a comment on the situation when the regularization matrix L is different from the identity. Then the minimization problem (10) is replaced by

$$\min_{\boldsymbol{x}\in\mathbb{K}_\ell(A,\boldsymbol{b})}\{\|A\boldsymbol{x}-\boldsymbol{b}\|^2+\frac{1}{\mu}\|L\boldsymbol{x}\|^2\},$$

which simplifies to

$$\min_{\boldsymbol{y}\in\mathbb{R}^\ell}\{\|H_{\ell+1,\ell}\boldsymbol{y}-\boldsymbol{e}_1\|\boldsymbol{b}\|\|^2+\frac{1}{\mu}\|R_\ell\boldsymbol{y}\|^2\}, \tag{26}$$

where $R_\ell\in\mathbb{R}^{\ell\times\ell}$ is the upper triangular matrix in the QR factorization of the matrix LV_ℓ. The reduced problem (26) may be solved by computing the generalized singular value decomposition of the matrix pair $\{H_{\ell+1,\ell},R_\ell\}$; more details are provided in [16]. The present paper focuses on the numerical aspects of the Arnoldi-Tikhonov method, e.g., on the loss of orthogonality of the columns of the computed matrix V_ℓ. It therefore suffices to consider the case when $L=I$.

3. Neubauer's method for determining the regularization parameter

Neubauer [20] discusses Tikhonov regularization of operator equations in infinite-dimensional Hilbert space. After discretization, a finite-dimensional equation of the form (1) is obtained. Neubauer [20] allows errors in both the matrix A and the right-hand side $\boldsymbol{b}$, and derives a cubic equation for the regularization parameter $\alpha=1/\mu$. We consider the special case of this equation when the matrix A is error-free. Let the matrices $V_{\ell+1}$ and $H_{\ell+1,\ell}$ be given by (8). Introduce the low-rank approximation

$$A^{(\ell)}=V_{\ell+1}H_{\ell+1,\ell}V_\ell^T$$

of A, and let $\mathcal{R}(A^{(\ell)})$ denote the range of $A^{(\ell)}$. Then the orthogonal projector $R_\ell:\mathbb{R}^n\to P_{\mathcal{R}(A^{(\ell)})}$ is given by

$$R_\ell=V_{\ell+1}U_{\ell+1}I_{q,\ell+1}U_{\ell+1}^TV_{\ell+1}^T,$$

where $I_{q,\ell+1}\in\mathbb{R}^{(\ell+1)\times(\ell+1)}$ is defined by

$$I_{q,\ell+1}=\begin{bmatrix} I_q & 0\\ 0 & 0\end{bmatrix}\in\mathbb{R}^{(\ell+1)\times(\ell+1)}$$

with I_q being the $q\times q$ identity matrix and $q \geq 0$ is the rank of the matrix $H_{\ell+1,\ell}$, and $U_{\ell+1}$ is the left orthogonal matrix in the SVD (18) of $H_{\ell+1,\ell}$. Neubauer's cubic equation for the regularization parameter $\alpha = 1/\mu > 0$ for the reduced problem (11) when A is error-free reads

$$\alpha^3 \boldsymbol{b}^T V_{\ell+1} U_{\ell+1} I_{q,\ell+1} \left(\Sigma_{\ell+1,\ell}\Sigma_{\ell+1,\ell}^T + \alpha I_{\ell+1}\right)^{-3} I_{q,\ell+1} U_{\ell+1}^T V_{\ell+1}^T \boldsymbol{b} = \delta^2, \quad (27)$$

where $\Sigma_{\ell+1,\ell}$ is the diagonal matrix in the SVD of $H_{\ell+1,\ell}$ (18), and δ is a bound for the error in $\boldsymbol{b}$ (cf. (5)); see [22] for further details about the situation when A is reduced by the Arnoldi process and is allowed to be contaminated by error. Neubauer [20, Proposition 3.1] shows that under suitable conditions, there is a unique $\alpha > 0$ that satisfies equation (27) and that the computed approximate solution of (4) with $L = I$ converges to the desired solution $\hat{\boldsymbol{x}}$ of (1) as the errors in A and $\boldsymbol{b}$ converge to zero. We are interested in comparing Neubauer's zero-finder to Newton's method in the situation when only the vector $\boldsymbol{b}$ is contaminated by error and investigating whether the former yields values of the regularization parameter that result in approximate solution of (1) that approximate the desired solution $\hat{\boldsymbol{x}}$ more accurately. We illustrate the performance of these zero-finders in Section 5. Applications of Neubauer's zero-finder to linear ill-posed problems that are reduced by the Arnoldi method and allow an error in the matrix A have recently been described in [22].

4. The flexible Arnoldi-Tikhonov method

For some matrices A and vectors $\boldsymbol{b}$, the Krylov subspace (9) determined as described in Section 2 might not contain vectors that represent known features of the desired solution $\hat{\boldsymbol{x}}$. For instance, it may be known that $\hat{\boldsymbol{x}}$ is a small perturbation of a vector with all entries equal, or a small perturbation of a vector, whose entries grow linearly as a function of their index. However, vectors in the generated Krylov subspace (9), which typically is of fairly small dimension ℓ, might not be able to represent such solutions accurately. In this situation one could try to increase the dimension ℓ until the Krylov subspace (9) contains vectors with desired properties, however, it often is cheaper to explicitly augment the generated Krylov subspace (9) by vectors that allow accurate approximation of $\hat{\boldsymbol{x}}$.

For definiteness, assume that we would like the vectors

$$\text{span}\{[1,1,\ldots,1]^T, [1,2,\ldots,n]^T\} \quad (28)$$

to live in the solution subspace. The Arnoldi-Tikhonov method (Algorithm 3) with the parameter $\ell_{\text{end}} = 0$ determines the Arnoldi decomposition

$$AV_{\ell_{\text{dis}}} = V_{\ell_{\text{dis}}+1} H_{\ell_{\text{dis}}+1,\ell_{\text{dis}}}. \quad (29)$$

We would like to enlarge the solution subspace $\mathcal{R}(V_{\ell_{\text{dis}}})$ by the vectors (28). This can be done by using the recursion formulas of the flexible Arnoldi method described by Saad [25]. The flexible GMRES method based on the flexible

Arnoldi method has previously been applied to the solution of linear discrete ill-posed problems in [19]. The novelty of the approach of this paper is to append the desired vectors (28) to an available Krylov subspace basis, instead of starting the construction of a solution subspace basis with the vectors (28). The approach used in [19] requires about twice the storage as the standard Arnoldi process to generate a solution subspace basis of the same dimension, while the approach of the present paper requires storage of only about $\ell_{\rm dis}+2\tilde{\ell}$ vectors in $\mathbb{R}^n$ to generate a solution subspace basis of dimension $\ell_{\rm dis}+\tilde{\ell}$, where $\tilde{\ell}$ is the number of non-Krylov vectors in the solution subspace. Typically, $\tilde{\ell}$ is chosen quite small.

We outline how to enlarge the solution subspace $\mathcal{R}(V_k)$ by including an arbitrary vector $\boldsymbol{u}\in\mathbb{R}^n$ that is not already in the subspace. Then the vector

$$\tilde{\boldsymbol{v}}_{k+1}=\frac{(I-V_kV_k^T)\boldsymbol{u}}{\|(I-V_kV_k^T)\boldsymbol{u}\|}$$

is orthogonal to the columns of the matrix $V_k=[\boldsymbol{v}_1,\boldsymbol{v}_2,\ldots,\boldsymbol{v}_k]$, and the columns of the matrix $\tilde{V}_{k+1}=[V_k,\tilde{\boldsymbol{v}}_{k+1}]$ form an orthonormal basis for the enlarged solution subspace. Let $\boldsymbol{w}:=A\tilde{\boldsymbol{v}}_{k+1}$, orthogonalize $\boldsymbol{w}$ against the columns of V_k, and normalize. The orthogonalization and normalization coefficients form column $\ell_{\rm dis}+1$ of the matrix $H_{\ell_{\rm dis}+2,\ell_{\rm dis}+1}\in\mathbb{R}^{\ell_{\rm dis}+2,\ell_{\rm dis}+1}$, whose leading $(\ell_{\rm dis}+1)\times\ell_{\rm dis}$ principal submatrix is the matrix $H_{\ell_{\rm dis}+1,\ell_{\rm dis}}$ in (29). Including $\tilde{\ell}-1$ additional vectors in the solution subspace gives the flexible Arnoldi decomposition

$$A\tilde{V}_{\ell_{\rm dis}+\tilde{\ell}}=V_{\ell_{\rm dis}+\tilde{\ell}+1}H_{\ell_{\rm dis}+\tilde{\ell}+1,\ell_{\rm dis}+\tilde{\ell}}. \tag{30}$$

Details about the computation of this decomposition are described in Algorithm 4. The approximate solution $\boldsymbol{x}_{\ell_{\rm dis}+\tilde{\ell}}$ of (1) can be expressed as

$$\boldsymbol{x}_{\ell_{\rm dis}+\tilde{\ell}}=\tilde{V}_{\ell_{\rm dis}+\tilde{\ell}}\boldsymbol{y}_{\ell_{\rm dis}+\tilde{\ell}},$$

where $\boldsymbol{y}_{\ell_{\rm dis}+\tilde{\ell}}\in\mathbb{R}^{\ell_{\rm dis}+\tilde{\ell}}$ solves the minimization problem

$$\min_{\boldsymbol{y}\in\mathbb{R}^{\ell_{\rm dis}+\tilde{\ell}}}\{\|H_{\ell_{\rm dis}+\tilde{\ell}+1,\ell_{\rm dis}+\tilde{\ell}}\boldsymbol{y}-\boldsymbol{e}_1\|\boldsymbol{b}\|\|^2+\frac{1}{\mu_{\ell_{\rm dis}+\tilde{\ell}}}\|\boldsymbol{y}\|^2\}.$$

Note that the first $\ell_{\rm dis}$ columns of the matrices $V_{\ell_{\rm dis}+\tilde{\ell}+1}$ and $\tilde{V}_{\ell_{\rm dis}+\tilde{\ell}}$ are the same and, therefore, can be stored in the same location.

Algorithm 4 (The Flexible Arnoldi Process).
Input: *matrix A, vector $\boldsymbol{b}\neq\boldsymbol{0}$, linearly independent vectors $\boldsymbol{u}_1,\boldsymbol{u}_2,\ldots,\boldsymbol{u}_{\tilde{\ell}}$ to be included in the solution subspace;*

1. *Compute the decomposition* (29) *with $V_{\ell_{\rm dis}}=[\boldsymbol{v}_1,\boldsymbol{v}_2,\ldots,\boldsymbol{v}_{\ell_{\rm dis}}]$, $V_{\ell_{\rm dis}+1}=[\boldsymbol{v}_1,\boldsymbol{v}_2,\ldots,\boldsymbol{v}_{\ell_{\rm dis}+1}]$, and $H_{\ell_{\rm dis}+1,\ell_{\rm dis}}=[h_{ij}]$ by Algorithm 3 with $\ell_{\rm end}=0$. The algorithm determines $\ell_{\rm dis}$.*
2. *Let $\tilde{V}_{\ell_{\rm dis}}=V_{\ell_{\rm dis}}$ to simplify notation. Actual storage space occupied by $\tilde{V}_{\ell_{\rm dis}}$ and $V_{\ell_{\rm dis}}$ may be the same;*

3. ***for*** $k = \ell_{\rm dis}, \ldots, \ell_{\rm dis} + \tilde{\ell} - 1$ ***do***
4. $\quad \tilde{\boldsymbol{v}}_{k+1} = \frac{(I - \tilde{V}_k \tilde{V}_k^T)\boldsymbol{u}_{k+1-\ell_{\rm dis}}}{\|(I - \tilde{V}_k \tilde{V}_k^T)\boldsymbol{u}_{k+1-\ell_{\rm dis}}\|}$;
5. $\quad \tilde{V}_{k+1} = [\tilde{V}_k, \tilde{\boldsymbol{v}}_{k+1}]$;
6. $\quad \boldsymbol{w} = A\tilde{\boldsymbol{v}}_{k+1}$;
7. $\quad$ ***for*** $j = 1, \ldots, k+1$ ***do***
8. $\quad\quad h_{j,k+1} = \boldsymbol{v}_j^T \boldsymbol{w}$;
9. $\quad\quad \boldsymbol{w} = \boldsymbol{w} - h_{j,k+1}\boldsymbol{v}_j$;
10. $\quad$ ***end for***
11. $\quad h_{k+2,k+1} = \|\boldsymbol{w}\|$;
12. $\quad \boldsymbol{v}_{k+2} = \boldsymbol{w}/h_{k+2,k+1}$;
13. $\quad V_{k+2} = [V_{k+1}, \boldsymbol{v}_{k+2}]$;
14. ***end for***
15. *Determine the regularization parameter* $\mu_{\ell_{\rm dis}+\tilde{\ell}}$ *by solving an analogue of* (21) *based on the decomposition* (30)*;*
16. *Solve the minimization problem*

$$\min_{\boldsymbol{y} \in \mathbb{R}^{\ell_{\rm dis}+\tilde{\ell}}} \{ \| H_{\ell_{\rm dis}+\tilde{\ell}+1, \ell_{\rm dis}+\tilde{\ell}} \boldsymbol{y} - \boldsymbol{e}_1 \|\boldsymbol{b}\| \|^2 + \frac{1}{\mu_{\ell_{\rm dis}+\tilde{\ell}}} \|\boldsymbol{y}\|^2 \}.$$

Denote the solution by $\boldsymbol{y}_{\ell_{\rm dis}+\tilde{\ell}}$*;*

17. $\boldsymbol{x}_{\ell_{\rm dis}+\tilde{\ell}} = \tilde{V}_{\ell_{\rm dis}+\tilde{\ell}} \boldsymbol{y}_{\ell_{\rm dis}+\tilde{\ell}}$*;*

Output*: ℓ_{dis}, the flexible Arnoldi decomposition* (30)*, and the approximate solution* $\boldsymbol{x}_{\ell_{\rm dis}+\tilde{\ell}}$ *of* (1)*.*

For ease of description of Algorithm 4, we assume that no breakdown occurs, i.e., that the vectors $\boldsymbol{u}_j$ are chosen so that $(I - \tilde{V}_k \tilde{V}_k^T)\boldsymbol{u}_{k+1-\ell_{\rm dis}} \neq \boldsymbol{0}$ in step 4 and $h_{k+1,k+1} \neq 0$ in step 12 of the algorithm. These restrictions can be worked around. They have not caused difficulties in the computed examples reported in the following section. The algorithm requires storage of about $\ell_{\rm dis} + 2\tilde{\ell}$ vectors in $\mathbb{R}^n$.

5. Computed examples

This section presents a few computed examples that illustrate the performance of the Arnoldi-Tikhonov (AT) method implemented by Algorithm 3 and of the flexible Arnoldi-Tikhonov (FAT) method implemented by Algorithm 4. The number of steps of these algorithms, as well as the regularization parameter, is determined by the discrepancy principle. We also will show that the AT method determines approximate solutions of higher quality than the Lanczos-Tikhonov (LT) method for the approximate solution of linear systems of equations (1) with a symmetric matrix. The LT method is obtained by replacing Algorithm 1 in Algorithm 3 by Algorithm 2. The regularization parameter μ is determined by Newton's method as described in Section 2, unless it is explicitly stated that the parameter is the solution of the cubic equation (27). The

matrices in all examples are of ill-determined rank. The error vector $\boldsymbol{e}$ in the right-hand side $\boldsymbol{b}$ in (1) simulates white Gaussian noise; thus, the entries of $\boldsymbol{e}$ are normally distributed pseudorandom numbers with mean zero and they are scaled to achieve a specified norm δ; cf. (5). We set $\eta = 1$ in all examples and let $\ell_{\text{init}} = 3$ in Algorithm 3. The desired solution $\hat{\boldsymbol{x}}$ of the error-free system (3) is known for all examples.

Let $\boldsymbol{x}_{\ell,\mu_\ell}$ denote the computed solution determined by one of the Krylov-Tikhonov methods. We measure the quality of computed solution by the relative error

$$\frac{\|\boldsymbol{x}_{\ell,\mu_\ell} - \hat{\boldsymbol{x}}\|}{\|\hat{\boldsymbol{x}}\|}.$$

The parameter ℓ_{dis} denotes the number of steps of the Arnoldi process required to satisfy the discrepancy principle. All computations are carried out using MATLAB with about 15 significant decimal digits.

Example 5.1. Regard the Fredholm integral equation of the first kind

$$\int_0^1 \kappa(s,t)x(t)dt = \exp(s) + (1-e)s - 1, \qquad 0 \leq s \leq 1, \tag{31}$$

with

$$\kappa(s,t) = \begin{cases} s(t-1), & s < t, \\ t(s-1), & s \geq t. \end{cases}$$

We discretize the integral equation by a Galerkin method with orthonormal box functions as test and trial functions using the MATLAB function deriv2 from [14]. This gives the symmetric matrix $A \in \mathbb{R}^{1000\times 1000}$. Its singular values decay to zero fairly slowly with increasing index number. Nevertheless, the matrix A is numerically singular. The function deriv2 also produces a scaled discrete approximation $\hat{\boldsymbol{x}} \in \mathbb{R}^{1000}$ of the solution $x(t) = \exp(t)$ of (31) and the error-free right-hand side vector $\hat{\boldsymbol{b}}$ in (3). An error vector $\boldsymbol{e}$ of norm δ is added to the vector $\hat{\boldsymbol{b}}$ to give the error-contaminated right-hand side vector $\boldsymbol{b}$; cf. (2).

Table 1 summarizes the computed results determined with Algorithm 3 when $\ell_{\text{end}} = 0$ and $\ell_{\text{end}} = 2$. We refer to this method as ATN when the regularization parameter is determined by Newton's method and by ATC when the regularization parameter is the solution of the cubic equation (27). The table illustrates that it is advantageous to carry out two more steps of the Arnoldi process after the discrepancy principle can be satisfied. Moreover, the table shows the relative error of the approximate solutions, when the regularization parameter is computed by Newton's method, to almost always be smaller than when the cubic zero finder is used to determine the regularization parameter; the error achieved with Newton's method is always smaller when $\ell_{\text{end}} = 2$. This suggests that determining the regularization parameter by Newton's method is beneficial. We found in this and other examples the cubic zero finder to determine values of the regularization parameter α that are larger than the values $1/\mu$ found by Newton's method; thus the cubic zero-finder gives approximate solutions of (1) that are more regularized than the approximate solutions determined with

Newton's method. Increasing $\ell_{\rm end}$ to 3 and 4 does not improve the quality of the computed solution; further $\ell_{\rm end} = 2$ gives solutions of higher quality than $\ell_{\rm end} = 1$.

δ	Method	$\ell_{\rm dis}$	$\frac{\|\boldsymbol{x}_{\ell_{\rm dis},\mu_{\ell_{\rm dis}}} - \hat{\boldsymbol{x}}\|}{\|\hat{\boldsymbol{x}}\|}$	$\frac{\|\boldsymbol{x}_{\ell_{\rm dis}+2,\mu_{\ell_{\rm dis}+2}} - \hat{\boldsymbol{x}}\|}{\|\hat{\boldsymbol{x}}\|}$
10^{-2}	ATN	3	$7.4203 \cdot 10^{-1}$	$3.2058 \cdot 10^{-1}$
	ATC	3	$4.9288 \cdot 10^{-1}$	$4.9037 \cdot 10^{-1}$
10^{-4}	ATN	9	$2.2788 \cdot 10^{-1}$	$1.8154 \cdot 10^{-1}$
	ATC	9	$2.0121 \cdot 10^{-1}$	$2.0032 \cdot 10^{-1}$
10^{-6}	ATN	22	$7.1578 \cdot 10^{-2}$	$7.0548 \cdot 10^{-2}$
	ATC	22	$7.7325 \cdot 10^{-2}$	$7.7263 \cdot 10^{-2}$

Table 1: Example 5.1: Results for different values of δ and $\ell_{\rm end}$. ATN stands for the Arnoldi-Tikhonov method with the regularization parameter determined by Newton's method, ATC denotes the Arnoldi-Tikhonov method with the regularization parameter determined by solving the cubic equation (27).

Results determined by Algorithm 3 with the Arnoldi process replaced by the Lanczos process are displayed in Table 2. This replacement is possible since the matrix A is symmetric. We refer to the method so defined as the LTN method. A comparison of Tables 1 and 2 shows the ATN method to furnish approximate solutions of higher quality than the LTN method. We also examined whether the Arnoldi process equipped with reorthogonalization gives computed solutions of higher quality than the Arnoldi process without reorthogonalization (as implemented by Algorithm 1), and found this not to be the case.

δ	Method	$\ell_{\rm dis}$	$\frac{\|\boldsymbol{x}_{\ell_{\rm dis},\mu_{\ell_{\rm dis}}} - \hat{\boldsymbol{x}}\|}{\|\hat{\boldsymbol{x}}\|}$	$\frac{\|\boldsymbol{x}_{\ell_{\rm dis}+2,\mu_{\ell_{\rm dis}+2}} - \hat{\boldsymbol{x}}\|}{\|\hat{\boldsymbol{x}}\|}$
10^{-2}	LTN	3	$5.7383 \cdot 10^{-1}$	$5.6829 \cdot 10^{-1}$
10^{-4}	LTN	3	$3.0165 \cdot 10^{-1}$	$2.8399 \cdot 10^{-1}$
10^{-6}	LTN	3	$3.0032 \cdot 10^{-1}$	$2.7633 \cdot 10^{-1}$

Table 2: Example 5.1: Results for Algorithm 3 with the Arnoldi process replaced by the Lanczos process. The regularization parameter is determined by Newton's method.

The solution of (31) is a monotonically increasing function. It therefore may be beneficial to include a basis for the subspace (28) into the solution subspace, because this subspace allows the representation of linearly increasing functions. Inclusion of the subspace (28) into the solution subspace means that the Krylov subspace part of the solution subspace only has to approximate the difference between the solution of (31) and a linear function. Table 3 illustrates that inclusion of the subspace (28) into the solution subspace, indeed, increases the quality of the computed solution. Specifically, comparing column four of

Table 1 with column four of Table 3 shows that enlarging the solution subspace $\mathbb{K}_{\ell_{\rm dis}}(A, \boldsymbol{b})$ with the subspace (28) increases the quality of the computed solution significantly.

The dimension of the solution subspace for Table 3 is $\ell_{\rm dis} + 2$, while the dimension of the solution subspace $\mathbb{K}_{\ell_{\rm dis}}(A, \boldsymbol{b})$ only is $\ell_{\rm dis}$. It is therefore also interesting to compare the errors of Table 3 to the error reported in the last column of Table 1. The errors in Table 3 are seen to be smaller than the errors in the last column of Table 1. We conclude that enlarging the solution subspace $\mathbb{K}_{\ell_{\rm dis}}(A, \boldsymbol{b})$ by the subspace (28) of dimension two gives an approximate solution of higher quality than when using the solution Krylov subspace $\mathbb{K}_{\ell_{\rm dis}+2}(A, \boldsymbol{b})$. We conclude from this that using an auxiliary subspace may give approximate solutions of higher quality than increasing the dimension of the Krylov subspace by the dimension of the auxiliary subspace.

			$\frac{\|\boldsymbol{x}_{\ell_{\rm dis}+\tilde{\ell},\mu_{\ell_{\rm dis}+\tilde{\ell}}} - \hat{\boldsymbol{x}}\|}{\|\hat{\boldsymbol{x}}\|}$
δ	Method	$\ell_{\rm dis}$	
10^{-2}	FAT	3	$3.0625 \cdot 10^{-1}$
10^{-4}	FAT	9	$1.0325 \cdot 10^{-1}$
10^{-6}	FAT	22	$3.9137 \cdot 10^{-2}$

Table 3: Example 5.1: Results for different values of δ for the FAT method with $\tilde{\ell} = 2$ and the auxiliary subspace (28). The regularization parameter is determined by Newton's method.

Example 5.2. Consider the Fredholm integral equation of the first kind

$$\int_{-\pi/2}^{\pi/2} \kappa(\tau, \sigma) x(\sigma) d\sigma = b(\tau), \qquad -\pi/2 \leq \tau \leq \pi/2, \tag{32}$$

with

$$\kappa(\sigma, \tau) = (\cos(\sigma) + \cos(\tau)) \left(\frac{\sin(\xi)}{\xi} \right)^2, \quad \xi = \pi(\sin(\sigma) + \sin(\tau)).$$

The right-hand side function $b(\tau)$ is chosen so that the solution $x(\sigma)$ is the sum of two Gaussian functions. This integral equation is discussed by Shaw [27]. We discretize it with the code shaw from [14], using a quadrature rule with 1000 nodes. This yields the symmetric matrix $A \in \mathbb{R}^{1000\times 1000}$ and the error-free right-hand side vector $\hat{\boldsymbol{b}}$. The matrix is numerically singular; its singular values decay to zero quite rapidly with increasing index number. The code shaw also produces a scaled discrete approximation $\hat{\boldsymbol{x}} \in \mathbb{R}^{1000}$ of the solution $x(\sigma)$ of (32). A noise-contaminated vector $\boldsymbol{b}$ is generated similarly as in Example 5.1.

The computed results are summarized in Table 4, which is analogous to Table 1. The table illustrates the advantage of adding two steps of the Arnoldi process after the discrepancy can be satisfied. The table shows that typically the computed solution is of higher quality when the regularization parameter is determined by Newton's method, than by solving the cubic equation (27).

δ	Method	$\ell_{\rm dis}$	$\frac{\|\boldsymbol{x}_{\ell_{\rm dis},\mu_{\ell_{\rm dis}}} - \hat{\boldsymbol{x}}\|}{\|\hat{\boldsymbol{x}}\|}$	$\frac{\|\boldsymbol{x}_{\ell_{\rm dis}+2,\mu_{\ell_{\rm dis}+2}} - \hat{\boldsymbol{x}}\|}{\|\hat{\boldsymbol{x}}\|}$
10^{-2}	ATN	9	$6.4457 \cdot 10^{-2}$	$3.3985 \cdot 10^{-2}$
	ATC	9	$5.3524 \cdot 10^{-2}$	$5.3501 \cdot 10^{-2}$
10^{-4}	ATN	10	$2.2449 \cdot 10^{-2}$	$2.0014 \cdot 10^{-2}$
	ATC	10	$3.2235 \cdot 10^{-2}$	$3.1980 \cdot 10^{-2}$
10^{-6}	ATN	12	$1.2523 \cdot 10^{-2}$	$1.1059 \cdot 10^{-2}$
	ATC	12	$1.6199 \cdot 10^{-2}$	$1.6118 \cdot 10^{-2}$

Table 4: Example 5.2: Results for different values of δ and $\ell_{\rm end}$. ATN stands for the Arnoldi-Tikhonov method with the regularization parameter determined by Newton's method, ATC denotes the Arnoldi-Tikhonov method with the regularization parameter determined by solving the cubic equation (27).

Table 5 complements Table 4 and shows the quality of the computed solutions when the Arnoldi process (Algorithm 1) is replaced by the Lanczos process (Algorithm 2) in Algorithm 3. A comparison of Tables 4 and 5 shows the Arnoldi process to give computed solutions of higher quality than the Lanczos process. Numerical experiments with an Arnoldi process that applies reorthogonalization did not increase the quality of the computed solutions.

δ	Method	ℓ_{dis}	$\frac{\|\boldsymbol{x}_{\ell_{dis},\mu_{\ell_{dis}}} - \hat{\boldsymbol{x}}\|}{\|\hat{\boldsymbol{x}}\|}$	$\frac{\|\boldsymbol{x}_{\ell_{dis}+2,\mu_{\ell_{dis}+2}} - \hat{\boldsymbol{x}}\|}{\|\hat{\boldsymbol{x}}\|}$
10^{-2}	LTN	3	$8.1272 \cdot 10^{-1}$	$7.9103 \cdot 10^{-1}$
10^{-4}	LTN	3	$8.1335 \cdot 10^{-1}$	$7.9156 \cdot 10^{-1}$
10^{-6}	LTN	3	$8.1336 \cdot 10^{-1}$	$7.9157 \cdot 10^{-1}$

Table 5: Example 5.2: Results for Algorithm 3 with the Arnoldi process replaced by the Lanczos process. The regularization parameter is determined by Newton's method.

Example 5.3. We turn to the Fredholm integral equation of the first kind discussed by Baart [1],

$$\int_0^{\frac{\pi}{2}} \kappa(s,t)x(t)dt = g(s), \qquad 0 \leq s \leq \pi, \tag{33}$$

where $\kappa(s,t) = \exp(s\cos(t))$ and $g(s) = 2\sinh(s)/s$. The solution is given by $x(t) = \sin(t)$. We discretize by a Galerkin method with 1000 orthonormal box functions as test and trial functions to obtain a nonsymmetric matrix $A \in \mathbb{R}^{1000\times 1000}$. Its singular values decrease to zero very quickly with increasing index number. The discretization is computed with the MATLAB function baart from [14]. The vector $\hat{\boldsymbol{x}} \in \mathbb{R}^{1000}$ is a discretization of the solution $x(t)$ of (33).

Table 6 summarizes the calculated results. The table is analogous to Table 1. The ATN approach can be seen to yield approximate solutions of higher quality

than the ATC method. We also note that using $\ell_{\rm end} = 2$ improves the quality of the solution compared with $\ell_{\rm end} = 0$. We found that reorthogonalization in the Arnoldi process does not improve the quality of the computed solution.

δ	Method	$\ell_{\rm dis}$	$\frac{\|\boldsymbol{x}_{\ell_{\rm dis},\mu_{\ell_{\rm dis}}} - \hat{\boldsymbol{x}}\|}{\|\hat{\boldsymbol{x}}\|}$	$\frac{\|\boldsymbol{x}_{\ell_{\rm dis}+2,\mu_{\ell_{\rm dis}+2}} - \hat{\boldsymbol{x}}\|}{\|\hat{\boldsymbol{x}}\|}$
10^{-2}	ATN	3	$1.0676 \cdot 10^{-1}$	$1.0293 \cdot 10^{-1}$
	ATC	3	$1.8813 \cdot 10^{-1}$	$2.2045 \cdot 10^{-1}$
10^{-5}	ATN	5	$4.5031 \cdot 10^{-2}$	$3.3954 \cdot 10^{-2}$
	ATC	5	$8.3205 \cdot 10^{-2}$	$6.7812 \cdot 10^{-2}$

Table 6: Example 5.3: Results for different values of δ and $\ell_{\rm end}$. ATN stands for the Arnoldi-Tikhonov method with the regularization parameter determined by Newton's method, ATC denotes the Arnoldi-Tikhonov method with the regularization parameter determined by solving the cubic equation (27).

Example 5.4. We would like to compute an approximate solution of the integral equation

$$\int_{-6}^{6} \kappa(s,t)x(t)dt = g(t), \qquad -6 \le t \le 6, \tag{34}$$

with kernel $\kappa(s,t) = f(s-t)$, where

$$f(t) = \begin{cases} 1 + \cos(\frac{\pi}{3}t), & \mid t \mid < 3, \\ 0, & \text{otherwise.} \end{cases}$$

and

$$g(s) = (6 - \mid s \mid)(1 + \frac{1}{2}\cos(\frac{\pi}{3}s)) + \frac{9}{2\pi}\sin(\frac{\pi}{3} \mid s \mid).$$

Integral equations of this form are discussed by Phillips [21]. We discretize the integral equation by a Nyström method based on a composite trapezoidal quadrature rule with 300 equidistant nodes. This gives the nonsymmetric matrix $A \in \mathbb{R}^{300\times 300}$, whose singular values converge to zero fairly slowly with increasing index number. Nevertheless, the matrix is numerically singular. The solution of (34) is $x(t) = f(t)$. Discretization of $x(s)$ at equidistant nodes gives the vector $\hat{\boldsymbol{x}} \in \mathbb{R}^{300}$. We let $\hat{\boldsymbol{b}} = A\hat{\boldsymbol{x}}$ and add an error vector $\boldsymbol{e}$ with normally distributed random entries to obtain the contaminated right-hand side $\boldsymbol{b}$ in (1).

Table 7 summarizes the calculated results and is analogous to Table 6. The ATN method can be seen to give approximate solutions of higher quality than the ATC approach. We also note that using $\ell_{\rm end} = 2$ improves the quality of the solution compared with $\ell_{\rm end} = 0$. We found that reorthogonalization in the Arnoldi process does not improve the quality of the computed solution.

6. Conclusion and extension

The performance of the Arnoldi-Tikhonov is investigated. We find that it is generally beneficial to carry out a few more steps of the Arnoldi process than

δ	Method	$\ell_{\rm dis}$	$\frac{\|\boldsymbol{x}_{\ell_{\rm dis},\mu_{\ell_{\rm dis}}} - \hat{\boldsymbol{x}}\|}{\|\hat{\boldsymbol{x}}\|}$	$\frac{\|\boldsymbol{x}_{\ell_{\rm dis}+2,\mu_{\ell_{\rm dis}+2}} - \hat{\boldsymbol{x}}\|}{\|\hat{\boldsymbol{x}}\|}$
10^{-2}	ATN	12	$4.3659 \cdot 10^{-3}$	$4.3069 \cdot 10^{-3}$
	ATC	12	$1.1580 \cdot 10^{-2}$	$1.1535 \cdot 10^{-2}$
10^{-4}	ATN	20	$8.2988 \cdot 10^{-4}$	$6.5825 \cdot 10^{-4}$
	ATC	20	$1.4757 \cdot 10^{-3}$	$1.4455 \cdot 10^{-3}$
10^{-6}	ATN	38	$1.0507 \cdot 10^{-4}$	$9.8722 \cdot 10^{-5}$
	ATC	38	$1.7636 \cdot 10^{-4}$	$1.7403 \cdot 10^{-4}$

Table 7: Example 5.4: Results for different values of δ and $\ell_{\rm end}$. ATN stands for the Arnoldi-Tikhonov method with the regularization parameter determined by Newton's method, ATC denotes the Arnoldi-Tikhonov method with the regularization parameter determined by solving the cubic equation (27).

necessary to satisfy the discrepancy principle. While it may be tempting to replace the Arnoldi process by the Lanczos process when the matrix A in (1) is symmetric, our numerical results suggest that this may not be a good idea due to loss of orthogonality of the Lanczos vectors, which results in computed solutions of inferior quality than solutions determined when the Arnoldi process is applied. We conclude that the Arnoldi process should be applied for both nonsymmetric and symmetric matrices A. Reorthogonalization in the Arnoldi process did not result in computed solutions of improved quality. A new algorithm for including auxiliary vectors in the solution subspace is described, and computed examples illustrate that it may be beneficial to enhance Krylov solution subspaces by user chosen-vectors. Finally, two approaches to determine the regularization parameter are compared. We found the approach based on Newton's method described in Section 2 to perform the best.

The computed examples in our comparison stem from the discretization of a few Fredholm integral equations of the first kind. The discretization gives rise to nonsymmetric or symmetric matrices with singular values that decay to zero quickly or slowly with increasing index number. The examples have been chosen to be representative of a variety of linear discrete ill-posed problems with a square system matrix.

The poor accuracy achieved when the Krylov subspace basis is determined by the Lanczos process raises the question whether the quality of computed solutions determined by applying Golub-Kahan bidiagonalization to A with initial vector $\boldsymbol{b}$ to compute a Krylov subspace basis can be improved by reorthogonalization. Golub-Kahan bidiagonalization uses a pair of short recurrence relations to generate the solution subspace; see, e.g., [2, 7] for details. This technique requires the evaluation of matrix-vector products with both the matrices A and A^T and therefore is outside the scope of methods discussed in this paper. Nevertheless, we carried out a few computed examples using Golub-Kahan bidiagonalization with and without reorthogonalization and found that the application of reorthogonalization typically does not result in computed solutions of significantly higher quality than when no reorthogonalization is carried out.

However, the discrepancy principle generally is satisfied after fewer iterations when reorthogonalization is applied. Numerical experiments showed that the number of Golub-Kahan bidiagonalization steps required to satisfy the discrepancy principle may be 50% larger without reorthogonalization.

Acknowledgment

M.A. would like to thank the Deanship of Scientific Research at Qassim University for funding his research for this project.

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), pp. 241–247.

[2] Å. Björck, A bidiagonalization algorithm for solving large and sparse ill-posed systems of linear equations, BIT Numer. Math., 18 (1988), pp. 659–670.

[3] A. Buccini, M. Pasha, and L. Reichel, Generalized singular value decomposition with iterated Tikhonov regularization, J. Comput. Appl. Math., 373 (2020), Art. 112276.

[4] A. Buccini and L. Reichel, An ℓ_p-ℓ_q minimization method with cross-validation for the restoration of impulse noise contaminated images, J. Comput. Appl. Math., 375 (2020), Art. 112824.

[5] D. Calvetti, B. Lewis, and L. Reichel, Krylov subspace iterative methods for nonsymmetric discrete ill-posed problems in image restoration, in Advanced Signal Processing Algorithms, Architectures, and Implementations XI, ed. F.T. Luk, Proceedings of the Society of Photo-Optical Instrumentation Engineers (SPIE), vol. 4474, The International Society for Optical Engineering, Bellingham, WA, 2001, pp. 224–233.

[6] D. Calvetti, S. Morigi, L. Reichel, and F. Sgallari, Tikhonov regularization and the L-curve for large, discrete ill-posed problems, J. Comput. Appl. Math., 123 (2000), pp. 423–446.

[7] D. Calvetti and L. Reichel, Tikhonov regularization of large linear problems, BIT Numer. Math., 43 (2003), pp. 263–283.

[8] T. F. Chan and K. R. Jackson, Nonlinearly preconditioned Krylov subspace methods for discrete Newton algorithms, SIAM J. Sci. Statist. Comput., 5 (1984), pp. 533–542.

[9] J. W. Daniel, W. B. Gragg, L. Kaufman, and G. W. Stewart, Reorthogonalization and stable algorithms for updating the Gram-Schmidt QR factorization, Math. Comp., 30 (1976), pp. 772–795.

[10] H. W. Engl, M. Hanke, and A. Neubauer, Regularization of Inverse Problems, Kluwer, Dordrecht, 1996.

[11] S. Gazzola and P. Novati, Multi-parameter Arnoldi-Tikhonov methods, Electron. Trans. Numer. Anal., 40 (2013), pp. 452–475.

[12] S. Gazzola, P. Novati, and M. R. Russo, On Krylov projection methods and Tikhonov regularization, Electron. Trans. Numer. Anal., 44 (2015), pp. 83–123.

[13] L. Greengard and V. Rokhlin, A new version of the fast multipole method for the Laplace equation in three dimensions, Acta Numerica, 6 (1997), pp. 229–269.

[14] P. C. Hansen, Regularization Tools, version 4.0 for Matlab 7.3, Numer. Algorithms, 46 (2007), pp. 189–194.

[15] P. C. Hansen, Rank-Deficient and Discrete Ill-Posed Problems, SIAM, Philadelphia, 1998.

[16] G. Huang, L. Reichel, and F. Yin, On the choice of solution subspace for nonstationary iterated Tikhonov regularization, Numer. Algorithms, 72 (2016), pp. 1043–1063.

[17] S. Kindermann, Discretization independent convergence rates for noise level-free parameter choice rules for the regularization of ill-conditioned problems, Electron. Trans. Numer. Anal., 38 (2011), pp. 233–257.

[18] S. Kindermann and K. Raik, A simplified L-curve method as error estimator, Electron. Trans. Numer. Anal., 53 (2020), pp. 217–238.

[19] K. Morikuni, L. Reichel, and K. Hayami, FGMRES for linear discrete ill-posed problems, Appl. Numer. Math., 75 (2014), pp. 175–187.

[20] A. Neubauer, An a posteriori parameter choice for Tikhonov regularization in the presence of modeling error, Appl. Numer. Math., 4 (1988), pp. 507–519.

[21] D. L. Phillips, A technique for the numerical solution of certain integral equations of the first kind, J. ACM, 9 (1962), pp. 84–97.

[22] R. Ramlau and L. Reichel, Error estimates for Arnoldi-Tikhonov regularization for ill-posed operator equations, Inverse Problems, 35 (2019), Art. 055002.

[23] L. Reichel and G. Rodriguez, Old and new parameter choice rules for discrete ill-posed problems, Numer. Algorithms, 63 (2013), pp. 65–87.

[24] L. Reichel and A. Shyshkov. A new zero-finder for Tikhonov regularization, BIT Numer. Math., 48 (2008), pp. 627–643.

[25] Y. Saad, A flexible inner-outer preconditioned GMRES algorithm, SIAM J. Sci. Comput., 14 (1993), pp. 461–469.

[26] Y. Saad, Iterative Methods for Sparse Linear Systems, 2nd ed., SIAM, Philadelphia, 2003.

[27] C. B. Shaw, Jr., Improvements of the resolution of an instrument by numerical solution of an integral equation, J. Math. Anal. Appl., 37 (1972), pp. 83–112.