Numerical Algorithms
https://doi.org/10.1007/s11075-025-02007-x

RESEARCH

Flexible iterative methods for linear systems of equations with multiple right-hand sides

Alessandro Buccini[1] · Marco Donatelli[2] · Lucas Onisk[3] · Lothar Reichel[4]

Received: 10 October 2024 / Accepted: 3 January 2025

Abstract
This paper describes new approaches to the solution of a sequence of large linear systems of equations or large linear least squares problems with the same matrix and several right-hand side vectors that represent data. We consider both the situations when the matrix of the systems to be solved is fairly well-conditioned and when the matrix is very ill-conditioned. In the latter case regularization is applied. We are concerned with the situation when the matrix is too large to make the application of direct solution methods possible or attractive. Our solution methods apply flexible Arnoldi or flexible Golub-Kahan decompositions. These decompositions allow the solution subspace computed during the solution of a seed system to be expanded by residual vectors that are computed during the solution of subsequent systems. Computed examples illustrate the competitiveness of the proposed methods.

Keywords Ill-posed problems · Iterative methods · Flexible Krylov subspaces · Tikhonov regularization

1 Introduction

We are concerned with computing solutions of a sequence of linear least-squares problems of the form

$$\min_{x^{(i)}\in\mathbb{R}^n} \left\| Ax^{(i)} - b^{(i)} \right\|, \qquad i = 1, 2, \ldots, k, \tag{1}$$

where $A \in \mathbb{R}^{m\times n}$, with $m \geq n$, is a large matrix and the vector $b^{(i)} \in \mathbb{R}^m$ represents data that may be contaminated by measurement errors, $e^{(i)} \in \mathbb{R}^m$, which we will refer to as "noise". Throughout this paper $\|\cdot\|$ denotes the Euclidean vector norm or the spectral matrix norm. When $m = n$, the least-squares problems (1) reduce to linear

Extended author information available on the last page of the article systems of equations; we will solve systems with $m > n$ and $m = n$ in the same manner.

When the matrix A is of small to moderate size, it is attractive to use a direct solution method that computes a factorization of A. We will assume that the matrix A is so large that the computation of its factorization is infeasible or unattractive and, therefore, focus on iterative solution methods. The methods described are most efficient when the data vectors $b^{(i)}$, $i = 1, 2, \ldots, k$, are fairly close, but this is not required for the applicability of the methods. For notational simplicity, we will assume that the vectors $b^{(i)}$ in (1) are scaled to be of unit Euclidean norm. This scaling affects the stopping criteria used for the iterative methods described in this paper. It is straightforward to allow arbitrary scalings of the $b^{(i)}$ when it is appropriate.

The solutions of a sequence of linear systems of equations with the same matrix A and different vectors $b^{(i)}$ has received considerable attention in the literature. This problem was first studied by Parlett [26] and Saad [30], who consider the situation when the matrix A is symmetric and positive definite. This case also is investigated by Abdel-Rehim et al. [1], Calvetti and Reichel [6], and Chan and Wan [7]. Simoncini and Gallopoulos [34] discuss the situation when the matrix A is square, nonsingular, and nonsymmetric.

If A stems from the discretization of an ill-posed problem, the matrix may be so ill-conditioned that Tikhonov regularization has to be used; see [12]. We consider both the situations when the matrix A is not very ill-conditioned and when A is severely ill-conditioned. Here the conditining is measured by the condition number, i.e., by the quotient of the largest and smallest singular values of A.

A popular solution approach for systems of equations (1) with a large symmetric symmetric positive definite, and not very ill-conditioned, matrix A is the application of a conjugate gradient-type method. Here the approximate solution of one of the systems, commonly referred to as the seed system, say $Ax^{(1)} = b^{(1)}$ is computed, and then approximate solutions of the remaining systems $Ax^{(i)} = b^{(i)}$, $i = 2, 3, \ldots, k$, are calculated by applying Galerkin projections into the Krylov subspace determined by the conjugate gradient method. The accuracy of the approximate solutions of the latter systems is, if necessary, enhanced by carrying out further iterations with the conjugate gradient method applied to these systems; see Abdel-Rehim et al. [1] for a discussion on computational aspects of this solution approach. Analogous solution methods can be used when A is a fairly general square nonsingular matrix; see Simoncini and Gallopoulos [34].

We present new approaches to the efficient solution of sequences of least-squares problems or linear systems of equations (1). Our approaches apply a flexible Arnoldi method or a flexible Golub-Kahan method to generate solution subspaces that contain residual vectors determined when when solving the problems (1). We believe the use of this kind of solution subspaces to be new. The first step of our methods for solving the problems (1) is to solve a seed system, say

$$\min_{x \in \mathbb{R}^n} \left\| Ax - b^{(1)} \right\|. \tag{2}$$

Denote the computed solution by $x^{(1)}$. We then seek to determine the solution of the next system,

$$\min_{x\in\mathbb{R}^n} \left\| Ax - b^{(2)} \right\|, \tag{3}$$

by using the solution subspace generated when solving the seed system and, if necessary, expand this solution subspace by vectors that represent residual errors for the system (3). Specifically, assume first that the matrix $A \in \mathbb{R}^{m\times n}$ is square and nonsingular, and that using the solution subspace determined when solving (2) does not yield a residual vector of small enough norm. This solution subspace then is enriched with the residual vector

$$r^{(2)} = b^{(2)} - Ax^{(1)}. \tag{4}$$

Enrichment is accomplished within the framework of the flexible GMRES method described by Saad [31]. Each step of the flexible GMRES method yields a new residual vector, which is included in the next solution subspace. This process is continued until an acceptable approximate solution $x^{(2)}$ of the system (3) is obtained. We then proceed to solve the system

$$\min_{x\in\mathbb{R}^n} \left\| Ax - b^{(3)} \right\| \tag{5}$$

by using the solution subspace determined for the solution of (3) and, if necessary, enrich the solution subspace with residual vectors for the system (5). The computations are continued in this manner until all k systems (1) have been solved. We remark that the solution subspace can be re-initialized, i.e., we discard the available solution subspace, if a basis for the solution subspace cannot be stored in the available computer storage. We have not encountered this issue in the computed examples reported in Section 6.

The flexible GMRES method is based on the flexible Arnoldi decomposition. If the matrix A is severely ill-conditioned, as it will be in some of applications that we consider, then we combine the flexible Arnoldi method with Tikhonov regularization.

When the matrix $A \in \mathbb{R}^{m\times n}$ in (2) is rectangular with $m > n$, we first compute an approximate solution $x^{(1)}$ of the least-squares problem (2) by a suitable number of steps of Golub-Kahan bidiagonalization applied to A with initial vector $b^{(1)}$, and then calculate an approximate solution of (3) by enriching the solution subspace determined when solving (2) by the residual

$$r^{(2)} = A^T b^{(2)} - A^T Ax^{(1)}$$

of the normal equations associated with the least-squares problem (3) in case enrichment is necessary; the superscript ^{T} denotes transposition. The enrichment of the solution subspace by the vector $r^{(2)}$ requires the application of a flexible Golub-Kahan decomposition method. Flexible Golub-Kahan decomposition methods have been described by Lampe et al. [21] and more recently by Chung and Gazzola [8]. We will use a slight modification of the latter method to solve the sequence of least-squares problems (1). The computations proceed analogously as in the situation when the matrix A is square, i.e., the least-squares problem (3) is solved by enriching the solution subspace with residual vectors for the normal equations associated with subsequently computed approximate solutions of (3) when a sufficiently accurate approximate solution cannot be found in the available solution subspace. Having computed a sufficiently accurate approximate solution $x^{(2)}$ of (3), we proceed to solve the least-squares problems (1) for $i = 3, 4, \ldots, k$ in a similar manner.

Our solution methods reuse and possibly expand solution subspaces that have been determined during the solution of the problems (1) for $i = 1, 2, \ldots, s-1$ when solving the problem $i = s$ for $s = 2, 3, \ldots, k$. Hence, these solution methods may be considered particular recycling methods. Recycling methods have received considerable attention in the literature; see, e.g., [2, 18, 35]. Our methods expand the solution subspaces by residual vectors. This expansion approach also is used in [4, 21, 22] for the solution of Tikhonov minimization problems in general form and nonlinear generalizations thereof.

The need to solve a sequence of systems (1) arises in a variety of applications. For instance, when seeking to solve a boundary value problem for a partial differential equation in three space-dimensions and the matrix A that represents the discretized differential operator requires a very large amount of computer storage, calculating the solution by solving a sequence of problems in two-space dimensions may be attractive. This situation is illustrated by Abdel-Rehim et al. [1]. Also, when solving problems of the form

$$\min_{X \in \mathbb{R}^{n \times k}} \|AX - B\|_F, \tag{6}$$

where $A \in \mathbb{R}^{m \times n}$, $m \geq n$, and $B = [b^{(1)}, b^{(2)}, \ldots, b^{(k)}] \in \mathbb{R}^{m \times k}$ with m, n, and k large, and $\| \cdot \|_F$ denotes the Frobenius norm, it may be attractive to compute each column of the solution matrix $X = [x^{(1)}, x^{(2)}, \ldots, x^{(k)}] \in \mathbb{R}^{n \times k}$ independently. This yields the least-squares problems (1) that require less computer storage to solve than the solution of (6) by an iterative method; see Simoncini and Gallopoulos [34] for a discussion of this situation.

We are interested in the situations when the matrix $A \in \mathbb{R}^{m \times n}$, $m \geq n$, is fairly well conditioned or when the matrix stems from the discretization of an ill-posed operator equation. The latter situation arises, for instance, when one seeks to restore video frames or some other sequence of related images that have been contaminated by blur and noise; see, e.g., Bentbib et al. [3] and Pasha et al. [27] for discussions and illustrations. In these applications often $m = n$, and k is the number of video frames or images. The matrix A represents a discretization of a blurring operator and, generally, is severely ill-conditioned; the data vectors $b^{(i)}$, $i = 1, 2, \ldots, k$, represent blur- and noise-contaminated images. We remark that the straightforward solution of the problems (6) by a truncated block GMRES method, where truncation yields a regularized approximate solution, may give computed approximate solutions of (1) of poor quality when the number of systems, k, is large, because the block Hessenberg matrix determined by the block GMRES method may be quite ill-conditioned. This can lead to numerical instability issues due to the noise in B; see [24] for a recent discussion. In this situation it is better to solve the systems (1) separately.

This paper is organized as follows. Section 2 describes our solution approach when the matrix A is square and fairly well conditioned. In Section 3, we consider the situation when A stems from the discretization of an operator with an unbounded inverse. Then A, typically, is severely ill-conditioned and, therefore, approximate solutions of

(1) are determined with the aid of Tikhonov regularization. Section 4 is concerned with the situation when the matrix $A \in \mathbb{R}^{m\times n}$ is rectangular, with $m \geq n$, and fairly well-conditioned. Each subproblem is solved by Golub-Kahan bidiagonalization applied to a seed system followed by flexible Golub-Kahan decomposition applied to the remaining problems. Section 5 describes the use of this solution technique complemented by Tikhonov regularization when the matrix A is very ill-conditioned. A few computed examples are presented in Section 6. Finally, Section 7 contains concluding remarks.

2 Solution of a sequence of linear systems of equations with a square fairly well-conditioned matrix by flexible Arnoldi decomposition

This section is concerned with the situation when the matrix $A \in \mathbb{R}^{n\times n}$ in (1) is fairly well-conditioned. The least-squares problems (1) reduce to linear systems of equations

$$Ax = b^{(i)}, \qquad i = 1, 2, \ldots, k.$$

We ignore possible "noise" in the right-hand side vectors in this section since the matrix is assumed to be fairly well-conditioned. The noise therefore will not propagate into the computed solution enough to significantly destroy the quality of the latter.

The first linear system, $Ax = b^{(1)}$, is referred to as the seed system. Application of E_1 steps of the Arnoldi process to A with initial vector $b^{(1)}$ gives the Arnoldi decomposition

$$AV_{E_1} = V_{E_1+1}H_{E_1+1,E_1}, \tag{7}$$

where the columns of the matrix $V_{E_1+1} \in \mathbb{R}^{n\times(E_1+1)}$ form an orthonormal basis for the Krylov subspace

$$\mathcal{K}_{E_1+1}(A, b^{(1)}) = \text{span}\left\{b^{(1)}, Ab^{(1)}, \ldots, A^{E_1}b^{(1)}\right\},$$

and $V_{E_1+1}e_1 = b^{(1)}/\|b^{(1)}\|$. Here and throughout this paper $e_1 = [1, 0, \ldots, 0]^T$ denotes the first principal axis vector. The matrix V_{E_1} is made up of the first E_1 columns of V_{E_1+1} and the matrix $H_{E_1+1,E_1} \in \mathbb{R}^{(E_1+1)\times E_1}$ is of upper Hessenberg form, i.e., all entries below the subdiagonal vanish. We tacitly assume that the decomposition (7) exists and that the matrix H_{E_1+1,E_1} is of full rank. This is the generic situation. Further details on the Arnoldi decomposition can be found in [32, Chapter 6], where also the GMRES method for the solution of the seed system is discussed. How to handle the (unusual) situation when the matrix H_{E_1+1,E_1} is rank-deficient is considered in [29].

Let $x = V_{E_1}y$ for $y \in \mathbb{R}^{E_1}$. In iteration E_1, the GMRES method applied to the seed system with initial approximate solution $x = 0$ solves the least-squares problem

$$\min_{x\in\mathcal{K}_{E_1}(A,b^{(1)})} \left\|Ax - b^{(1)}\right\| = \min_{y\in\mathbb{R}^{E_1}} \left\|H_{E_1+1,E_1}y - \|b_1\|e_1\right\|, \tag{8}$$

where the right-hand side is obtained by substituting the decomposition (7) into the left-hand side; see, e.g., [32, Chapter 6] for further details. Denote the approximate solution of the small least-squares problem on the right-hand side of (8) by $y^{(1)} \in \mathbb{R}^{E_1}$. This gives the approximate solution $x^{(1)} = V_{E_1} y^{(1)}$ of the large least-squares problem on the left-hand side. We choose the number of iterations E_1 so that $x^{(1)}$ satisfies

$$\left\| Ax^{(1)} - b^{(1)} \right\| = \left\| H_{E_1+1,E_1} y^{(1)} - \|b^{(1)}\| e_1 \right\| \leq \delta, \tag{9}$$

where $\delta > 0$ is a user-chosen tolerance.

Consider for the moment the Arnoldi decompositions

$$AV_\ell = V_{\ell+1} H_{\ell+1,\ell}, \qquad \ell = 1, 2, \ldots, E_1,$$

where the matrix V_ℓ consists of the first ℓ columns of V_{E_1+1} and the matrix $H_{\ell+1,\ell}$ is the leading $(\ell + 1) \times \ell$ principal submatrix of H_{E_1+1,E_1}. The columns of V_ℓ form an orthonormal basis for the Krylov subspace $\mathcal{K}_\ell(A, b^{(1)})$. In iteration ℓ, for $\ell = 1, 2, \ldots, E_1$, GMRES solves the least-squares problem

$$\min_{x \in \mathcal{K}_\ell(A, b^{(1)})} \left\| Ax - b^{(1)} \right\| = \min_{y \in \mathbb{R}^\ell} \left\| H_{\ell+1,\ell} y - \|b^{(1)}\| e_1 \right\| \tag{10}$$

by computing the solution $y_\ell^{(1)}$ of the minimization problem on the right-hand side and then determining the solution $x_\ell^{(1)} = V_\ell y_\ell^{(1)}$ of the left-hand side problem, in a similar manner as equation (8) is solved.

Theorem 1 *Define the residual vectors*

$$r_\ell^{(1)} = b^{(1)} - Ax_\ell^{(1)}, \qquad \ell = 1, 2, \ldots, E_1.$$

These vectors belong to the Krylov subspace $\mathcal{K}_{E_1+1}\left(A, b^{(1)}\right)$.

Proof For every $\ell = 1, 2, \ldots, E_1$, the vector $x_\ell^{(1)}$ lives in the Krylov subspace $\mathcal{K}_\ell\left(A, b^{(1)}\right)$. Hence, the vector $r_\ell^{(1)} - b^{(1)}$ is in the subspace $A\mathcal{K}_\ell\left(A, b^{(1)}\right)$ and, therefore, the vector $r_\ell^{(1)}$ belongs to the Krylov subspace $\mathcal{K}_{\ell+1}\left(A, b^{(1)}\right)$. □

The above theorem shows that the solution subspace for GMRES when solving (8) is made up of the residual vectors determined by GMRES when solving the systems (10) for $\ell = 1, 2, \ldots, E_1$. This suggests that the span of residual vectors also may be a good choice of solution subspace when the span is not a Krylov subspace. The methods in this paper use this kind of solution subspaces and computed examples presented in Section 6 illustrate that spans of residual vectors indeed make up suitable solution subspaces.

After having solved the seed system (9) by GMRES to desired accuracy, we have the approximate solution $x_{E_1}^{(1)}$, which we henceforth will denote by $x^{(1)}$. We turn to the solution of the linear system of equations

$$Ax = b^{(2)}. \tag{11}$$

If the vector $b^{(2)}$ is very close to $b^{(1)}$, then an accurate approximate solution may be available in the subspace $\mathcal{K}_{E_1}(A, b^{(1)})$ (which is indicated by the fact the the inequality (13) below hold). We therefore solve

$$\min_{y \in \mathbb{R}^{E_1}} \left\| A V_{E_1} y - b^{(2)} \right\| . \tag{12}$$

The solution of (12) can be evaluated by solving the small least-squares problem

$$\min_{y \in \mathbb{R}^{E_1}} \left\| H_{E_1+1,E_1} y - V_{E_1+1}^T b^{(2)} \right\| .$$

Denote the solution by $y_{E_1}^{(2)}$. Then $x_{E_1}^{(2)} = V_{E_1} y_{E_1}^{(2)}$ is an approximate solution of (12). Assume first that the inequality

$$\left\| A x_{E_1}^{(2)} - b^{(2)} \right\| \leq \delta \tag{13}$$

holds. Then we accept $x_{E_1}^{(2)}$ as a solution of (11), which we henceforth denote by $x^{(2)}$, and continue on to the linear system of equations

$$Ax = b^{(3)}.$$

We remark that in order for $x^{(2)}$ to be an accurate approximation of the solution of (3), it is generally required that the columns of the matrix V_{E_1+1} be numerically orthogonal. Therefore we implement Arnoldi's method with re-orthogonalization.

We turn to the situation when inequality (13) does not hold. Then we compute the residual vector (4). This vector is orthogonalized against the available solution subspace and normalized to give the vector

$$v = \frac{r^{(2)} - V_{E_1} V_{E_1}^T r^{(2)}}{\|r^{(2)} - V_{E_1} V_{E_1}^T r^{(2)}\|} \in \mathbb{R}^n,$$

where we assume that the vector $r^{(2)} - V_{E_1} V_{E_1}^T r^{(2)}$ does not vanish. This is the generic situation. The solution subspace range(V_{E_1}) is enriched with the vector v. Define the matrix

$$\widetilde{V}_{E_1+1} = [V_{E_1}, v] \in \mathbb{R}^{n \times (E_1+1)}.$$

Its orthonormal columns span the new solution subspace. The recursion formulas for the flexible Arnoldi process yield an upper Hessenberg matrix $\widetilde{H}_{E_1+2,E_1+1} \in \mathbb{R}^{(E_1+2)\times(E_1+1)}$ and a matrix $\widetilde{U}_{E_1+2} \in \mathbb{R}^{n \times (E_1+2)}$ with orthonormal columns such that

$$A \widetilde{V}_{E_1+1} = \widetilde{U}_{E_1+2} \widetilde{H}_{E_1+2,E_1+1}; \tag{14}$$

see Saad [31] for details, where also the closely related flexible GMRES method is described. For the convenience of the reader, we provide Algorithm 1 that describes the flexible Arnoldi process. It is implemented with re-orthogonalization.

Algorithm 1 Flexible Arnoldi process.

Input: $A \in \mathbb{R}^{n\times n}$, $b^{(j)} \in \mathbb{R}^n$ for $1 \leq j \leq k$, number of steps s
Output: $\widetilde{V}_k \in \mathbb{R}^{n\times k}$, $\widetilde{H}_{k+1,k} \in \mathbb{R}^{(k+1)\times k}$, and $\widetilde{U}_{k+1} \in \mathbb{R}^{m\times(j+1)}$

Set $\widetilde{u}_1 = b^{(j)}/\|b^{(j)}\|$ for some $j \in \{1, 2, \ldots, k\}$
for $k = 1, 2, \ldots, s$ **do**
  Specify $\widetilde{v}_k$
  $q = A\widetilde{v}_k$
  for $i = 1, 2, \ldots, k$ **do**
    $\widetilde{h}_{i,k} = q^T\widetilde{u}_i$ and $q = q - \widetilde{h}_{i,k}\widetilde{u}_i$
  end
  $\widetilde{h}_{k+1,k} = \|q\|$ and $\widetilde{u}_{k+1} = q/\widetilde{h}_{k+1,k}$
end

Algorithm 1 allows the vectors $\widetilde{v}_k$ to be fairly arbitrary. This requires storage of the matrix $\widetilde{U}_{k+1}$. However, in our applications of the algorithm many of these vectors are columns of the seed matrix V_{E_1}. Therefore, many of the columns of $\widetilde{U}_{k+1}$ agree with columns of V_{E_1} and, hence, do not have to be explicitly stored. For simplicity of exposition, we ignore this aspect in the description of Algorithm 1. Moreover, the vectors $\widetilde{v}_k$ either are provided as input or computed during the execution of the algorithm. The algorithm as described carries out s steps and only picks one data vector $b^{(j)}$.

Using the decomposition (14), the flexible GMRES method solves the minimization problem

$$\min_{x\in\text{span}(\widetilde{V}_{E_1+1})} \left\| Ax - b^{(2)} \right\| \tag{15}$$

with solution $x^{(2)}_{E_1+1}$. This solution can be evaluated by first solving the small minimization problem

$$\min_{y\in\mathbb{R}^{E_1+1}} \left\| \widetilde{H}_{E_1+2,E_1+1}y - \widetilde{U}^T_{E_1+2}b^{(2)} \right\| \tag{16}$$

with solution y. Then $x^{(2)}_{E_1+1} = \widetilde{V}_{E_1+1}y$. This follows from (14). When exploiting the relation between the minimization problems (15) and (16) in finite-precision computation, it is important that the columns of $\widetilde{U}_{E_1+2}$ be numerically orthonormal. We therefore implement the flexible Arnoldi process with re-orthogonalization.

Introduce the residual vector

$$r^{(2)}_{E_1+1} = b^{(2)} - Ax^{(2)}_{E_1+1}$$

and append it, after orthogonalization to range$(\widetilde{V}_{E_1+1})$ and normalization, to the matrix $\widetilde{V}_{E_1+1}$ to give the matrix $\widetilde{V}_{E_1+2}$. The flexible GMRES method now gives a new approximate solution to (11). We continue in this manner to enlarge the solution subspace until we have determined a decomposition

$$A\widetilde{V}_{E_2} = \widetilde{U}_{E_2+1}\widetilde{H}_{E_2+1,E_2}, \tag{17}$$

where the matrices $\widetilde{V}_{E_2} \in \mathbb{R}^{n\times E_2}$ and $\widetilde{U}_{E_2+1} \in \mathbb{R}^{n\times(E_2+1)}$ have orthonormal columns, $\widetilde{H}_{E_2+1,E_2} \in \mathbb{R}^{(E_2+1)\times E_2}$ is of upper Hessenberg form, and

$$\min_{x\in\text{span}(\widetilde{V}_{E_2})} \left\| Ax - b^{(2)} \right\| \leq \delta.$$

We now turn to the solution of (5) and proceed analogously as described above when solving (3). The decomposition (17) is expanded for every system of equations to be solved. When all systems (1) have been solved, we have the decomposition

$$A\widetilde{V}_{E_k} = \widetilde{U}_{E_k+1}\widetilde{H}_{E_k+1,E_k}, \tag{18}$$

for some integer $E_k > 0$. The details of the computations are described by Algorithm 2.

We remark that to keep the amount of computer storage needed bounded, the solution subspace can be re-initialized when its dimension is larger than a user-specified parameter dim_{max}. We will not dwell on this issue because re-initialization was not required in the computed examples described in Section 6.

3 Solution of a sequence of linear systems of equations with a square severely ill-conditioned matrix by flexible Arnoldi decomposition

In this section the matrix $A \in \mathbb{R}^{n\times n}$ is severely ill-conditioned. We therefore apply Tikhonov regularization to determine approximate solutions of the linear systems of equations (1). Tikhonov regularization replaces the solution of these systems with the solution of regularized least-squares problems

$$\min_{x\in\mathbb{R}^n} \left\{ \left\| Ax - b^{(i)} \right\|^2 + \mu^{(i)} \left\| L^{(i)}x \right\|^2 \right\}, \qquad i = 1, 2, \ldots, k, \tag{19}$$

where the $L^{(i)} \in \mathbb{R}^{s\times n}$ are regularization matrices and the $\mu^{(i)} > 0$ are regularization parameters. The number of rows s may be larger, smaller, or equal to n. We will assume that the matrices $L^{(i)}$ are chosen so that null spaces of A and $L^{(i)}$ intersect trivially for all i. Then the minimization problems (19) have unique solutions $x^{(i)}$ for every $\mu^{(i)} > 0$ and $i = 1, 2, \ldots, k$. In many applications, the matrices $L^{(i)}$ are chosen to be discretizations of differential operators; see [12, 17] for discussions and illustrations.

The data vectors $b^{(i)}$ are assumed to be contaminated by noise. Since the matrix A is very ill-conditioned, the noise in $b^{(i)}$ has to be taken into account during the solution process. Straightforward solution of the problems (1) generally does not give meaningful computed approximate solutions due to severe propagation of the errors $e^{(i)}$ in the $b^{(i)}$ into the computed solutions. Let $b^{(i)}_{true}$ denote the unknown noise-free vector associated with $b^{(i)}$, i.e.,

$$b^{(i)} = b^{(i)}_{true} + e^{(i)}, \quad i = 1, 2, \ldots, k. \tag{20}$$

Algorithm 2 Flexible Arnoldi for well-conditioned linear systems.

Input: $A \in \mathbb{R}^{n\times n}$, $b^{(j)},\ j = 1, 2, \ldots k$, and tol
Output: $x^{(j)} \in \mathbb{R}^{n},\ j = 1, 2, \ldots, k$

for $i = 1, 2, \ldots, n$ **do**
  Carry out one step of Arnoldi: $AV_i = V_{i+1}H_{i+1,i}$, with $V_1 = b^{(1)}/\|b^{(1)}\|$
  Solve $\hat{y} = \arg\min_{y\in\mathbb{R}^i} \left\| H_{i+1,i}\, y - \|b^{(1)}\| e_1 \right\|$
  Compute $r_i^{(1)} = b^{(1)} - AV_i\,\hat{y}$
  if $\|r_i^{(1)}\| < tol$ **then**
    $x^{(1)} = V_i\,\hat{y}$
    $E_2 = i$
  end
end
for $j = 2, \ldots, k$ **do**
  $u = 1$
  Solve $\hat{y} = \arg\min_{y\in\mathbb{R}^{E_j}} \left\| \widetilde{H}_{E_j+u,E_j}\, y - \widetilde{U}^T_{E_j+u} b^{(j)} \right\|$
  Compute $r_u^{(j)} = b^{(j)} - A\widetilde{V}_{E_j}\,\hat{y}$
  if $\|r_u^{(j)}\| < tol$ **then**
    $x^{(j)} = \widetilde{V}_{E_j}\,\hat{y}$
    $E_{j+1} = E_j$
  else
    $v = \left(r_u^{(j)} - \widetilde{V}_{E_j}\widetilde{V}^T_{E_j} r_u^{(j)}\right) / \left\| r_u^{(j)} - \widetilde{V}_{E_j}\widetilde{V}^T_{E_j} r_u^{(j)} \right\|$
    $u = u + 1$
    $\widetilde{V}_{E_j+u} = \left[\widetilde{V}_{E_j+u-1},\ v\right]$
    Carry out one step of flexible Arnoldi: $A\widetilde{V}_{E_j+u} = \widetilde{U}_{E_j+u+1}\widetilde{H}_{E_j+u+1,E_j+u}$
    Solve $\hat{y} = \arg\min_{y\in\mathbb{R}^{E_j+u}} \left\| \widetilde{H}_{E_j+u+1,E_j+u}\, y - \widetilde{U}^T_{E_j+u+1} b^{(j)} \right\|$
    Compute $r_u^{(j)} = b^{(j)} - A\widetilde{V}_{E_j+u}\,\hat{y}$
    while $\|r_u^{(j)}\| \geq tol$ **do**
      $v = \left(r_u^{(j)} - \widetilde{V}_{E_j+u}\widetilde{V}^T_{E_j+u} r_u^{(j)}\right) / \left\| r_u^{(j)} - \widetilde{V}_{E_j+u}\widetilde{V}^T_{E_j+u} r_u^{(j)} \right\|$
      $u = u + 1$
      $\widetilde{V}_{E_j+u} = \left[\widetilde{V}_{E_j+u-1},\ v\right]$
      Carry out one step of flexible Arnoldi: $A\widetilde{V}_{E_j+u} = \widetilde{U}_{E_j+u+1}\widetilde{H}_{E_j+u+1,E_j+u}$
      Solve $\hat{y} = \arg\min_{y\in\mathbb{R}^{E_j+u}} \left\| \widetilde{H}_{E_j+u+1,E_j+u}\, y - \widetilde{U}^T_{E_j+u+1} b^{(j)} \right\|$
      Compute $r_u^{(j)} = b^{(j)} - A\widetilde{V}_{E_j+u}\,\hat{y}$
    end
    $x^{(j)} = \widetilde{V}_{E_j+u}\,\hat{y}$
    $E_{j+1} = E_j + u$
  end
end

Let error bounds

$$\left\| e^{(i)} \right\| \leq \delta, \qquad i = 1, 2, \ldots, k, \tag{21}$$

be known. We remark that this value of δ does not have to be the same as in (9). It is straightforward to allow each error $e^{(i)}$ to have a different bound $\delta^{(i)} > 0$. Below we will comment on the situation no bound for the errors $e^{(i)}$ is available.

We are interested in determining accurate approximations of the solutions of the unavailable noise-free problems associated with the known noise-contaminated discrete ill-posed problems (1). Thus, we would like to solve the unavailable least-squares problems

$$\min_{x \in \mathbb{R}^n} \left\| Ax - b_{true}^{(i)} \right\|, \qquad i = 1, 2, \ldots, k,$$

and denote the solutions of minimal Euclidean norm by $x_{true}^{(i)}$, $i = 1, 2, \ldots, k$. The regularization parameter $\mu^{(i)}$ determines how sensitive the solution $x^{(i)}$ of (19) is to the error $e^{(i)}$ in $b^{(i)}$, and how close $x^{(i)}$ is to the desired solution $x_{true}^{(i)}$. A small value of $\mu^{(i)}$ makes the solution of (19) more sensitive to the error $e^{(i)}$ than a large value, but an unnecessarily large value often results that the solution may lack details the exact solution $x_{true}^{(i)}$ may possess.

The discrepancy principle dictates that the regularization parameter $\mu^{(i)} > 0$ be chosen so that

$$\left\| Ax^{(i)} - b^{(i)} \right\| = \tau\delta, \qquad i = 1, 2, \ldots, k, \tag{22}$$

where $\tau \geq 1$ is a user-specified parameter. We remark that if a bound δ in (21) is not known, then so-called heuristic approaches can be used to determine a suitable value of δ. These approaches include the L-curve criterion, cross validation, and generalized cross validation; see, e.g., [12, 19, 20, 28] for discussions and illustrations of the performance of heuristic methods.

The solution subspaces are determined similarly as in Section 2. Let the matrix $\widetilde{V}_{E_k}$ be the same as in the decomposition (18). We replace the Tikhonov minimization problems (19) by the low-dimensional minimization problems

$$\min_{y \in \mathbb{R}^{E_k}} \left\{ \left\| A\widetilde{V}_{E_k} y - b^{(i)} \right\|^2 + \mu^{(i)} \left\| L^{(i)} \widetilde{V}_{E_k} y \right\|^2 \right\}, \qquad i = 1, 2, \ldots, k. \tag{23}$$

These minimization problems can be simplified by using the decomposition (18) and the QR factorizations

$$L^{(i)} \widetilde{V}_{E_k} = Q_{E_k}^{(i)} R_{E_k}^{(i)}, \qquad i = 1, 2, \ldots, k,$$

where the matrix $Q_{E_k}^{(i)} \in \mathbb{R}^{n \times E_k}$ has orthonormal columns and $R_{E_k}^{(i)} \in \mathbb{R}^{E_k \times E_k}$ is upper triangular. The following small minimization problems

$$\min_{y \in \mathbb{R}^{E_k}} \left\{ \left\| H_{E_k+1,E_k} y - \widetilde{U}_{E_k+1}^T b^{(i)} \right\|^2 + \mu^{(i)} \left\| R_{E_k}^{(i)} y \right\|^2 \right\}, \qquad i = 1, 2, \ldots, k, \tag{24}$$

are equivalent to the problems (23). Let $y_{E_k}^{(i)} \in \mathbb{R}^{E_k}$ denote the solution of the ith minimization problem (24) with $\mu^{(i)}$ determined by the discrepancy principle; see below. The vectors $y_{E_k}^{(i)}$ may, for example, be calculated by using the generalized singular value decomposition of the matrix pair $\{H_{E_k+1,E_k}, R_{E_k}^{(i)}\}$; see [10, 12]. Cheaper alternatives are described in [9, 17]. Then $x_{E_k}^{(i)} = \widetilde{V}_{E_k} y_{E_k}^{(i)}$ is an approximate solution of the ith Tikhonov minimization problem (23).

The discrepancy principle (22) can be evaluated according to

$$\begin{aligned}\left\| A x_{E_k}^{(i)} - b^{(i)} \right\|^2 &= \left\| A \widetilde{V}_{E_k} y_{E_k}^{(i)} - b^{(i)} \right\|^2 = \left\| \widetilde{U}_{E_k+1} \widetilde{H}_{E_k+1,E_k} y_{E_k}^{(i)} - b^{(i)} \right\|^2 \\ &= \left\| \widetilde{H}_{E_k+1,E_k} y_{E_k}^{(i)} - \widetilde{U}_{E_k+1}^T b^{(i)} \right\|^2 + \left\| b^{(i)} - \widetilde{U}_{E_k+1} \widetilde{U}_{E_k+1}^T b^{(i)} \right\|^2 \end{aligned} \tag{25}$$

The last term in (25) generally is very close to zero. In the computations reported in Section 6, we ignore the last term and determine the regularization parameter $\mu^{(i)}$ so that

$$\left\| \widetilde{H}_{E_k+1,E_k} y_{E_k}^{(i)} - \widetilde{U}_{E_k+1}^T b^{(i)} \right\| = \tau \delta.$$

Details of the computations are described by Algorithm 3.

4 Solution of a sequence of least-squares problems with a fairly well-conditioned matrix by flexible Golub-Kahan decomposition

This section differs from Section 2 in that the matrix $A \in \mathbb{R}^{m \times n}$ in (1) is allowed to be rectangular with $m \geq n$, and the Arnoldi and flexible Arnoldi decompositions are replaced by Golub-Kahan bidiagonalization and flexible Golub-Kahan decomposition, respectively. We outline the solution method but omit some details that can be inferred from the discussion in Section 2.

Similarly as above, we refer to the least-squares problem in (1) with index $i = 1$ as the seed system. Application of E_1 steps of Golub-Kahan bidiagonalization to A with initial vector $b^{(1)}$ gives the Golub-Kahan decompositions

$$A V_{E_1} = U_{E_1+1} B_{E_1+1,E_1}, \qquad A^T U_{E_1} = V_{E_1} B_{E_1,E_1}^T, \tag{26}$$

where the columns of the matrices $U_{E_1+1} \in \mathbb{R}^{m \times (E_1+1)}$ and $V_{E_1} \in \mathbb{R}^{n \times E_1}$ are orthonormal with $U_{E_1+1} e_1 = b^{(1)}/\|b^{(1)}\|$. The matrix $U_{E_1} \in \mathbb{R}^{m \times E_1}$ is made up of the first E_1 columns of U_{E_1+1}. Moreover, the columns of V_{E_1} form a basis for the Krylov subspace

$$\mathcal{K}_{E_1}\left(A^T A, A^T b^{(1)}\right) = \text{span}\left\{A^T b^{(1)}, (A^T A) A^T b^{(1)}, \ldots, (A^T A)^{E_1 - 1} A^T b^{(1)}\right\}.$$

The matrix $B_{E_1+1,E_1} \in \mathbb{R}^{(E_1+1) \times E_1}$ is lower bidiagonal with positive subdiagonal entries and B_{E_1,E_1} denotes its leading $E_1 \times E_1$ submatrix. We assume that the decompositions (26) with the stated properties exists. This is the generic situation.

Algorithm 3 Regularized flexible Arnoldi for ill-conditioned linear systems.

Input: $A \in \mathbb{R}^{n\times n}$, $b^{(j)} \in \mathbb{R}^n$, $L^{(j)} \in \mathbb{R}^{s\times n}$, $j = 1, 2, \ldots k$, τ and δ

Output: $x^{(j)} \in \mathbb{R}^n$, $j = 1, 2, \ldots, k$

for $i = 1, 2, \ldots, n$ **do**

  Carry out one step of Arnoldi: $AV_i = V_{i+1}H_{i+1,i}$, with $V_1 = b^{(1)}/\|b^{(1)}\|$

  Compute QR factorization: $\left[Q^{(1)}, R^{(1)}\right] = L^{(1)}V_i$

  Solve $\hat{y} = \arg\min_{y\in\mathbb{R}^i}\left\{\left\|\begin{bmatrix} H_{i+1,i} \\ \mu^{(1)}R^{(1)} \end{bmatrix} y - \begin{bmatrix} \|b^{(1)}\|e_1 \\ 0 \end{bmatrix}\right\|\right\}$ s.t $\left\|H_{i+1,i}\hat{y} - \|b^{(1)}\|e_1\right\|^2 = \tau^2\delta^2$

  Compute $r_i^{(1)} = b^{(1)} - AV_i\hat{y}$

  if $\|r_i^{(1)}\| \le \tau\delta$ **then**

    $x^{(1)} = V_i\hat{y}$

    $E_2 = i$

  end

end

for $j = 2, \ldots, k$ **do**

  $u = 1$

  Compute QR factorization: $\left[Q^{(j)}, R^{(j)}\right] = L^{(j)}\widetilde{V}_{E_j}$

  Solve $\hat{y} = \arg\min_{y\in\mathbb{R}^{E_j}}\left\{\left\|\begin{bmatrix} \widetilde{H}_{E_j+u,E_j} \\ \mu^{(j)}R^{(j)} \end{bmatrix} y - \begin{bmatrix} \widetilde{U}^T_{E_j+u}b^{(j)} \\ 0 \end{bmatrix}\right\|\right\}$ s.t $\left\|\widetilde{H}_{E_j+u,E_j}\hat{y} - \widetilde{U}^T_{E_j+u}b^{(j)}\right\|^2 = \tau^2\delta^2$

  Compute $r_u^{(j)} = b^{(j)} - A\widetilde{V}_{E_j}\hat{y}$

  if $\|r_u^{(j)}\| < \tau\delta$ **then**

    $x^{(j)} = \widetilde{V}_{E_j}\hat{y}$

    $E_{j+1} = E_j$

  else

    $v = \left(r_u^{(j)} - \widetilde{V}_{E_j}\widetilde{V}^T_{E_j}r_u^{(j)}\right) / \left\|r_u^{(j)} - \widetilde{V}_{E_j}\widetilde{V}^T_{E_j}r_u^{(j)}\right\|$

    $u = u + 1$

    $\widetilde{V}_{E_j+u} = \left[\widetilde{V}_{E_j+u-1},\ v\right]$

    Carry out one step of flexible Arnoldi: $A\widetilde{V}_{E_j+u} = \widetilde{U}_{E_j+u+1}\widetilde{H}_{E_j+u+1,E_j+u}$

    Compute QR factorization: $\left[Q^{(j)}, R^{(j)}\right] = L^{(j)}\widetilde{V}_{E_j+u}$

    Solve $\hat{y} = \arg\min_{y\in\mathbb{R}^{E_j+u}}\left\{\left\|\begin{bmatrix} \widetilde{H}_{E_j+u+1,E_j+u} \\ \mu^{(j)}R^{(j)} \end{bmatrix} y - \begin{bmatrix} \widetilde{U}^T_{E_j+u+1}b^{(j)} \\ 0 \end{bmatrix}\right\|\right\}$ s.t $\left\|\widetilde{H}_{E_j+u+1,E_j+u}\hat{y} - \widetilde{U}^T_{E_j+u+1}b^{(j)}\right\|^2 = \tau^2\delta^2$

    Compute $r_u^{(j)} = b^{(j)} - A\widetilde{V}_{E_j+u}\hat{y}$

    while $\|r_u^{(j)}\| \ge \tau\delta$ **do**

      $v = \left(r_u^{(j)} - \widetilde{V}_{E_j+u}\widetilde{V}^T_{E_j+u}r_u^{(j)}\right) / \left\|r_u^{(j)} - \widetilde{V}_{E_j+u}\widetilde{V}^T_{E_j+u}r_u^{(j)}\right\|$

      $u = u + 1$

      $\widetilde{V}_{E_j+u} = \left[\widetilde{V}_{E_j+u-1},\ v\right]$

      Carry out one step of flexible Arnoldi: $A\widetilde{V}_{E_j+u} = \widetilde{U}_{E_j+u+1}\widetilde{H}_{E_j+u+1,E_j+u}$

      Compute QR factorization: $\left[Q^{(j)}, R^{(j)}\right] = L^{(j)}\widetilde{V}_{E_j+u}$

      Solve $\hat{y} = \arg\min_{y\in\mathbb{R}^{E_j+u}}\left\{\left\|\begin{bmatrix} \widetilde{H}_{E_j+u+1,E_j+u} \\ \mu^{(j)}R^{(j)} \end{bmatrix} y - \begin{bmatrix} \widetilde{U}^T_{E_j+u+1}b^{(j)} \\ 0 \end{bmatrix}\right\|\right\}$ s.t $\left\|\widetilde{H}_{E_j+u+1,E_j+u}\hat{y} - \widetilde{U}^T_{E_j+u+1}b^{(j)}\right\|^2 = \tau^2\delta^2$

      Compute $r_u^{(j)} = b^{(j)} - A\widetilde{V}_{E_j+u}\hat{y}$

    end

    $x^{(j)} = \widetilde{V}_{E_j+u}\hat{y}$

    $E_{j+1} = E_j + u$

  end

end

Let $x = V_{E_1} y$ for $y \in \mathbb{R}^{E_1}$. In iteration E_1, the left-hand side decomposition (26) yields

$$\min_{x \in \mathcal{K}_{E_1}(A^T A, A^T b^{(1)})} \left\| Ax - b^{(1)} \right\| = \min_{y \in \mathbb{R}^{E_1}} \left\| B_{E_1+1,E_1} y - \|b^{(1)}\| e_1 \right\|. \tag{27}$$

Denote the solution of the small least-squares problem on the right-hand side by $y_{E_1} \in \mathbb{R}^{E_1}$. This gives the approximate solution $x_{E_1}^{(1)} = V_{E_1} y_{E_1}$ of the large least-squares problem on the left-hand side of (27). We choose the number of steps E_1 so that $x_{E_1}^{(1)}$ satisfies

$$\left\| A x_{E_1}^{(1)} - b^{(1)} \right\| \leq \tau \delta. \tag{28}$$

For notational simplicity, we will refer to the solution $x_{E_1}^{(1)}$ of (28) as $x^{(1)}$.

Consider the Golub-Kahan decompositions

$$A V_\ell = U_{\ell+1} B_{\ell+1,\ell}, \qquad A^T U_\ell = V_\ell B_{\ell,\ell}, \qquad \ell = 1, 2, \ldots, E_1, \tag{29}$$

where the matrix V_ℓ consists of the first ℓ columns of V_{E_1}, the matrices $U_{\ell+1}$ and U_ℓ are made up of the first $\ell + 1$ and ℓ columns of U_{E_1+1}, respectively, and $B_{\ell+1,\ell}$ is the leading $(\ell + 1) \times \ell$ submatrix of B_{E_1+1,E_1}. The columns of V_ℓ form an orthonormal basis for the Krylov subspace $\mathcal{K}_\ell(A^T A, A^T b^{(1)})$. In iteration ℓ, for $\ell = 1, 2, \ldots, E_1$, the decompositions (29) can be applied to solve the least-squares problems

$$\min_{x \in \mathcal{K}_\ell(A, b^{(1)})} \left\| Ax - b^{(1)} \right\| = \min_{y \in \mathbb{R}^\ell} \left\| B_{\ell+1,\ell} y - \|b^{(1)}\| e_1 \right\|, \quad \ell = 1, 2, \ldots, E_1, \tag{30}$$

by computing the solutions $y_\ell^{(1)}$ of the small problems on the right-hand sides and then determining the approximate solutions $x_\ell = V_\ell y_\ell^{(1)}$ of the left-hand side problems in a similar manner as equation (27) is solved. The proof of the following result is analogous to the proof of Theorem 1.

Theorem 2 *Define the solution vectors $x_\ell^{(1)}$ and the residual vectors for the normal equations associated with the left-hand side of* (30)

$$r_\ell^{(1)} = A^T b^{(1)} - A^T A x_\ell^{(1)}, \qquad \ell = 1, 2, \ldots, E_1.$$

These vectors belong to the Krylov subspace $\mathcal{K}_{E_1+1}\left(A^T A, A^T b^{(1)}\right)$.

Similarly as in Section 2, the above theorem suggests that the span of residual vectors may be a good choice of solution subspace also when they do not span a Krylov subspace.

After having solved the seed system (28) to desired accuracy, we have the approximate solution $x^{(1)}$ and turn to the solution of the next least-squares problem, i.e., of the problem (1) with index $i = 2$. A sufficiently accurate approximate solution of this system may live in the available Krylov subspace $\mathcal{K}_{E_1}(A^TA, A^Tb^{(1)})$. We therefore solve

$$\min_{y\in\mathbb{R}^{E_1}} \left\| AV_{E_1}y - b^{(2)} \right\|. \tag{31}$$

Note that the product AV_{E_1} already is known; see (26). The solution of (31) can be determined by solving the small least-squares problem

$$\min_{y\in\mathbb{R}^{E_1}} \left\| B_{E_1+1,E_1}y - U_{E_1+1}^T b^{(2)} \right\|.$$

Denote the solution by $y_{E_1}^{(2)}$. Then $x_{E_1}^{(2)} = V_{E_1}y_{E_1}^{(2)}$ is a solution of (31). If

$$\left\| Ax_{E_1}^{(2)} - b^{(2)} \right\| \leq \tau\delta, \tag{32}$$

then we accept $x_{E_1}^{(2)}$ as a solution of (31) and move on to compute a solution of the next least-squares problem in the sequence (1). Note that, analogously to Section 2, the columns of the matrix U_{E_1+1} should be numerically orthogonal. We therefore implement Golub-Kahan bidiagonalization with re-orthogonalization.

If inequality (32) does not hold, then we compute the residual vector

$$r_{E_1}^{(2)} = A^Tb^{(2)} - A^TAx_{E_1}^{(2)}, \tag{33}$$

which can be evaluated quite inexpensively as

$$r_{E_1}^{(2)} = A^T\left(b^{(2)} - U_{E_1+1}B_{E_1+1,E_1}y_{E_1}^{(2)}\right).$$

Proposition 1 *In exact arithmetic, the vector* (33) *is orthogonal to the Krylov subspace* $\mathcal{K}_{E_1}(A^TA, A^Tb^{(2)})$.

Proof This follows from the fact that the vector $y_{E_1}^{(2)}$ solves (31). We have $V_{E_1}^TA^TAV_{E_1}y_{E_1}^{(2)} = V_{E_1}^TA^Tb^{(2)}$. Substituting $x_{E_1}^{(2)} = V_{E_1}y_{E_1}^{(2)}$ into (33) shows that $V_{E_1}^Tr_{E_1}^{(2)} = 0$. □

Since we are using computer arithmetic, we re-orthogonalize $r_{E_1}^{(2)}$ against the available solution subspace and normalize this vector to obtain

$$\widetilde{v}_{E_1+1} = \frac{r_{E_1}^{(2)} - V_{E_1}V_{E_1}^Tr_{E_1}^{(2)}}{\|r_{E_1}^{(2)} - V_{E_1}V_{E_1}^Tr_{E_1}^{(2)}\|}.$$

We assume that the vector $r_{E_1}^{(2)} - V_{E_1}V_{E_1}^Tr_{E_1}^{(2)}$ does not vanish. This is the generic situation. The solution subspace range(V_{E_1}) is enriched by the vector $\widetilde{v}_{E_1+1}$. Introduce the matrix

$$\widetilde{V}_{E_1+1} = [V_{E_1}, \widetilde{v}_{E_1+1}] \in \mathbb{R}^{n\times(E_1+1)}.$$

Using a slight modification of the flexible Golub-Kahan process described by Chung and Gazzola [8], we obtain the decompositions

$$A\widetilde{V}_{E_1+1} = \widetilde{U}_{E_1+2}\widetilde{M}_{E_1+1}, \qquad A^T\widetilde{U}_{E_1+1} = \widetilde{W}_{E_1+1}\widetilde{R}^T_{E_1+1}. \tag{34}$$

The matrices of this decomposition are of the same sizes as the analogous matrices obtained after E_1+1 steps of standard Golub-Kahan bidiagonalization. Thus, $\widetilde{U}_{E_1+2} \in \mathbb{R}^{m\times(E_1+2)}$ has orthonormal columns with the matrix $\widetilde{U}_{E_1+1} = U_{E_1+1}$ from (26) its leading $m \times (E_1+1)$ submatrix, and the matrix $\widetilde{M}_{E_1+1} \in \mathbb{R}^{(E_1+2)\times(E_1+1)}$ is of upper Hessenberg form with leading $(E_1+1)\times E_1$ submatrix B_{E_1+1,E_1}. The matrix $\widetilde{W}_{E_1+1}$ agrees with U_{E_1+1} in (26) and $\widetilde{R}_{E_1+1} \in \mathbb{R}^{(E_1+1)\times(E_1+1)}$ is upper triangular with leading $E_1 \times E_1$ submatrix $B^T_{E_1,E_1}$.

We would like to compute the solution of the large minimization problem

$$\min_{x\in\mathrm{range}(\widetilde{V}_{E_1+1})} \left\| Ax - b^{(2)} \right\| = \min_{y\in\mathbb{R}^{E_1+1}} \left\| A\widetilde{V}_{E_1+1}y - b^{(2)} \right\|, \tag{35}$$

which is equivalent to the small minimization problem

$$\min_{y\in\mathbb{R}^{E_1+1}} \left\| \widetilde{M}_{E_1+1}y - \widetilde{U}^T_{E_1+2}b^{(2)} \right\|.$$

The solution $y^{(2)}_{E_1+1}$ of the latter problem gives the solution $x^{(2)}_{E_1+1} = \widetilde{V}_{E_1+1}y^{(2)}_{E_1+1}$ of (35). We compute the associated residual vector

$$r^{(2)}_{E_1+1} = A^Tb^{(2)} - A^TAx^{(2)}_{E_1+1}, \tag{36}$$

which can be evaluated inexpensively as

$$r^{(2)}_{E_1+1} = A^Tb^{(2)} - A^T\widetilde{U}_{E_1+2}\widetilde{M}_{E_1+1}y^{(2)}_{E_1+1},$$

where we note that the vector $A^Tb^{(2)}$ already has been calculated above. The following result follows similarly as Proposition 1.

Proposition 2 *In exact arithmetic, the vector* (36) *is orthogonal to the solution subspace* $\mathrm{range}(\widetilde{V}_{E_1+1})$.

The residual vector $r^{(2)}_{E_1+1}$ is used to enrich the solution subspace $\widetilde{V}_{E_1+1}$ similarly as we expanded the solution subspace with the previous residual vector above. Thus, let

$$\widetilde{v}_{E_1+2} = \frac{r^{(2)}_{E_1+1} - V_{E_1+1}V^T_{E_1+1}r^{(2)}_{E_1+1}}{\|r^{(2)}_{E_1+1} - V_{E_1+1}V^T_{E_1+1}r^{(2)}_{E_1+1}\|},$$

where we assume that $r^{(2)}_{E_1+1} - V_{E_1+1}V^T_{E_1+1}r^{(2)}_{E_1+1} \neq 0$. Introduce the matrix

$$\widetilde{V}_{E_1+2} = [V_{E_1+1}, \widetilde{v}_{E_1+2}] \in \mathbb{R}^{n\times(E_1+2)}.$$

Analogously to (34), we define the decompositions

$$A\widetilde{V}_{E_1+2} = \widetilde{U}_{E_1+3}\widetilde{M}_{E_1+2}, \qquad A^T\widetilde{U}_{E_1+2} = \widetilde{W}_{E_1+2}\widetilde{R}^T_{E_1+2}, \tag{37}$$

where $\widetilde{U}_{E_1+3} \in \mathbb{R}^{m\times(E_1+3)}$ has orthonormal columns with leading $m \times (E_1 + 2)$ submatrix $\widetilde{U}_{E_1+2}$, and the matrix $\widetilde{M}_{E_1+2} \in \mathbb{R}^{(E_1+3)\times(E_1+3)}$ is of upper Hessenberg form with leading $(E_1 + 2) \times (E_1 + 1)$ submatrix M_{E_1+2,E_1+1}. Finally, the matrix $\widetilde{W}_{E_1+2}$ has leading principal submatrix $\widetilde{W}_{E_1+1}$ and $\widetilde{R}_{E_1+2} \in \mathbb{R}^{(E_1+2)\times(E_1+2)}$ is upper triangular with leading submatrix $R^T_{E_1+1,E_1+1}$. The details of the computations are described by Algorithm 4.

We conclude this section by noting that the decompositions on the right in (34) and (37) are not needed for the computations, only the decompositions on the left-hand sides are used. The latter decompositions are determined by the matrices A and $\widetilde{V}_{E_1+\ell}$, which define $\widetilde{U}_{E_1+\ell+1}$ and $\widetilde{M}_{E_1+\ell+1}$. The decompositions on the left-hand sides are QR factorizations of $A^T\widetilde{U}_{E_1+\ell+1}$. These observations lead directly to the following result.

Theorem 3 *The decompositions*

$$A\widetilde{V}_{E_1+\ell} = \widetilde{U}_{E_1+\ell+1}\widetilde{M}_{E_1+\ell+1}, \qquad \ell = 1, 2, \dots ,$$

are flexible Arnoldi decompositions with the leading submatrices in the decomposition determined by Golub-Kahan bidiagonalization.

5 Solution of a sequence of least-squares problems with a severely ill-conditioned matrix by flexible Golub-Kahan decomposition

The solution method of this section differs from the one described in Section 3 only in that the flexible Arnoldi method is replaced by the flexible Golub-Kahan method. Tikhonov regularization is applied similarly as in Section 3. We omit a detailed discussion of the method, but provide Algorithm 5 which shows the computations. This algorithm is used for the computed examples in Section 6.

6 Computed examples

We illustrate the performance of the flexible Arnoldi (FA) and flexible Golub-Kahan (FGK) methods, as well as their regularized variants, the regularized flexible Arnoldi (RFA) and regularized flexible Golub-Kahan (RFGK) methods, when applied to the solution of sequences of well-posed or ill-posed problems. We compare our non-regularized methods to benchmark methods such as GMRES and a truncated Golub-Kahan (truncated GK) method which is mathematically equivalent to LSQR. The regularized variants, i.e., RFA and RFGK are compared to what we term 'Arnoldi-Tikhonov' and 'Golub-Kahan Tikhonov' (GK-Tikhonov). We note regarding the latter method that this is the most appropriate comparison instead of, e.g., LSQR, since all

Algorithm 4 Flexible Golub-Kahan for well-conditioned linear systems.

Input: $A \in \mathbb{R}^{m\times n}$, $b^{(j)} \in \mathbb{R}^m$, $j = 1, 2, \ldots k$, and tol
Output: $x^{(j)} \in \mathbb{R}^n$, $j = 1, 2, \ldots, k$
for $i = 1, 2, \ldots, n$ **do**
  Carry out one step of GK: $AV_i = U_{i+1}B_{i+1,i}$, with $U_1 = b^{(1)}/\|b^{(1)}\|$
  Solve $\hat{y} = \arg\min_{y\in\mathbb{R}^i} \left\| B_{i+1,i}\, y - \|b^{(1)}\| e_1 \right\|$
  Compute $r_i^{(1)} = b^{(1)} - AV_i\,\hat{y}$
  if $\|r_i^{(1)}\| < tol$ **then**
    $x^{(1)} = V_i\,\hat{y}$
    $E_2 = i$
  end
end
for $j = 2, \ldots, k$ **do**
  $u = 1$
  Solve $\hat{y} = \arg\min_{y\in\mathbb{R}^{E_j}} \left\| \widetilde{M}_{E_j+u,E_j}\, y - \widetilde{U}_{E_j+u}^T b^{(j)} \right\|$
  Compute $r_u^{(j)} = b^{(j)} - A\widetilde{V}_{E_j}\,\hat{y}$
  if $\|r_u^{(j)}\| < tol$ **then**
    $x^{(j)} = \widetilde{V}_{E_j}\,\hat{y}$
    $E_{j+1} = E_j$
  else
    $v = \left(A^T r_u^{(j)} - \widetilde{V}_{E_j}\widetilde{V}_{E_j}^T A^T r_u^{(j)}\right) / \left\| A^T r_u^{(j)} - \widetilde{V}_{E_j}\widetilde{V}_{E_j}^T A^T r_u^{(j)} \right\|$
    $u = u + 1$
    $\widetilde{V}_{E_j+u} = \left[\widetilde{V}_{E_j+u-1},\ v\right]$
    Carry out one step of flexible GK: $A\widetilde{V}_{E_j+u} = \widetilde{U}_{E_j+u+1}\widetilde{M}_{E_j+u+1,E_j+u}$
    Solve $\hat{y} = \arg\min_{y\in\mathbb{R}^{E_j+u}} \left\| \widetilde{M}_{E_j+u+1,E_j+u}\, y - \widetilde{U}_{E_j+u+1}^T b^{(j)} \right\|$
    Compute $r_u^{(j)} = b^{(j)} - A\widetilde{V}_{E_j+u}\,\hat{y}$
    while $\|r_u^{(j)}\| \geq tol$ **do**
      $v = \left(A^T r_u^{(j)} - \widetilde{V}_{E_j+u}\widetilde{V}_{E_j+u}^T A^T r_u^{(j)}\right) / \left\| A^T r_u^{(j)} - \widetilde{V}_{E_j+u}\widetilde{V}_{E_j+u}^T A^T r_u^{(j)} \right\|$
      $u = u + 1$
      $\widetilde{V}_{E_j+u} = \left[\widetilde{V}_{E_j+u-1},\ v\right]$
      Carry out one step of flexible GK: $A\widetilde{V}_{E_j+u} = \widetilde{U}_{E_j+u+1}\widetilde{M}_{E_j+u+1,E_j+u}$
      Solve $\hat{y} = \arg\min_{y\in\mathbb{R}^{E_j+u}} \left\| \widetilde{M}_{E_j+u+1,E_j+u}\, y - \widetilde{U}_{E_j+u+1}^T b^{(j)} \right\|$
      Compute $r_u^{(j)} = b^{(j)} - A\widetilde{V}_{E_j+u}\,\hat{y}$
    end
    $x^{(j)} = \widetilde{V}_{E_j+u}\,\hat{y}$
    $E_{j+1} = E_j + u$
  end
end

Algorithm 5 Regularized flexible Golub-Kahan for ill-conditioned linear systems.

Input: $A \in \mathbb{R}^{m\times n}$, $b^{(j)} \in \mathbb{R}^m$, $L^{(j)} \in \mathbb{R}^{s\times n}$, $j = 1, 2, \dots k$, τ and δ

Output: $x^{(j)} \in \mathbb{R}^n$, $j = 1, 2, \dots, k$

for $i = 1, 2, \dots, n$ **do**
- Carry out one step of GK: $AV_i = U_{i+1}B_{i+1,i}$, with $U_1 = b^{(1)}/\|b^{(1)}\|$
- Compute QR factorization: $\left[Q^{(1)}, R^{(1)}\right] = L^{(1)}V_i$
- Solve $\hat{y} = \arg\min_{y\in\mathbb{R}^i} \left\{ \left\| \begin{bmatrix} B_{i+1,i} \\ \mu^{(1)}R^{(1)} \end{bmatrix} y - \begin{bmatrix} \|b^{(1)}\|e_1 \\ 0 \end{bmatrix} \right\| \right\}$ s.t $\left\| B_{i+1,i}\hat{y} - \|b^{(1)}\|e_1 \right\|^2 = \tau^2\delta^2$
- Compute $r_i^{(1)} = b^{(1)} - AV_i\hat{y}$
- **if** $\|r_i^{(1)}\| \le \tau\delta$ **then**
  - $x^{(1)} = V_i\hat{y}$
  - $E_2 = i$
- **end**

end

for $j = 2, \dots, k$ **do**
- $u = 1$
- Compute QR factorization: $\left[Q^{(j)}, R^{(j)}\right] = L^{(j)}\widetilde{V}_{E_j}$
- Solve $\hat{y} = \arg\min_{y\in\mathbb{R}^{E_j}} \left\{ \left\| \begin{bmatrix} \widetilde{M}_{E_j+u,E_j} \\ \mu^{(j)}R^{(j)} \end{bmatrix} y - \begin{bmatrix} \widetilde{U}^T_{E_j+u}b^{(j)} \\ 0 \end{bmatrix} \right\| \right\}$ s.t $\left\| \widetilde{M}_{E_j+u,E_j}\hat{y} - \widetilde{U}^T_{E_j+u}b^{(j)} \right\|^2 = \tau^2\delta^2$
- Compute $r_u^{(j)} = b^{(j)} - A\widetilde{V}_{E_j}\hat{y}$
- **if** $\|r_u^{(j)}\| < \tau\delta$ **then**
  - $x^{(j)} = \widetilde{V}_{E_j}\hat{y}$
  - $E_{j+1} = E_j$
- **else**
  - $v = \left(A^Tr_u^{(j)} - \widetilde{V}_{E_j}\widetilde{V}^T_{E_j}A^Tr_u^{(j)}\right) / \left\| A^Tr_u^{(j)} - \widetilde{V}_{E_j}\widetilde{V}^T_{E_j}A^Tr_u^{(j)} \right\|$
  - $u = u + 1$
  - $\widetilde{V}_{E_j+u} = \left[\widetilde{V}_{E_j+u-1}, v\right]$
  - Carry out one step of flexible GK: $A\widetilde{V}_{E_j+u} = \widetilde{U}_{E_j+u+1}\widetilde{M}_{E_j+u+1,E_j+u}$
  - Compute QR factorization: $\left[Q^{(j)}, R^{(j)}\right] = L^{(j)}\widetilde{V}_{E_j+u}$
  - Solve $\hat{y} = \arg\min_{y\in\mathbb{R}^{E_j+u}} \left\{ \left\| \begin{bmatrix} \widetilde{M}_{E_j+u+1,E_j+u} \\ \mu^{(j)}R^{(j)} \end{bmatrix} y - \begin{bmatrix} \widetilde{U}^T_{E_j+u+1}b^{(j)} \\ 0 \end{bmatrix} \right\| \right\}$ s.t $\left\| \widetilde{M}_{E_j+u+1,E_j+u}\hat{y} - \widetilde{U}^T_{E_j+u+1}b^{(j)} \right\|^2 = \tau^2\delta^2$
  - Compute $r_u^{(j)} = b^{(j)} - A\widetilde{V}_{E_j+u}\hat{y}$
  - **while** $\|r_u^{(j)}\| \ge \tau\delta$ **do**
    - $v = \left(A^Tr_u^{(j)} - \widetilde{V}_{E_j+u}\widetilde{V}^T_{E_j+u}A^Tr_u^{(j)}\right) / \left\| A^Tr_u^{(j)} - \widetilde{V}_{E_j+u}\widetilde{V}^T_{E_j+u}A^Tr_u^{(j)} \right\|$
    - $u = u + 1$
    - $\widetilde{V}_{E_j+u} = \left[\widetilde{V}_{E_j+u-1}, v\right]$
    - Carry out one step of flexible GK: $A\widetilde{V}_{E_j+u} = \widetilde{U}_{E_j+u+1}\widetilde{M}_{E_j+u+1,E_j+u}$
    - Compute QR factorization: $\left[Q^{(j)}, R^{(j)}\right] = L^{(j)}\widetilde{V}_{E_j+u}$
    - Solve $\hat{y} = \arg\min_{y\in\mathbb{R}^{E_j+u}} \left\{ \left\| \begin{bmatrix} \widetilde{M}_{E_j+u+1,E_j+u} \\ \mu^{(j)}R^{(j)} \end{bmatrix} y - \begin{bmatrix} \widetilde{U}^T_{E_j+u+1}b^{(j)} \\ 0 \end{bmatrix} \right\| \right\}$ s.t $\left\| \widetilde{M}_{E_j+u+1,E_j+u}\hat{y} - \widetilde{U}^T_{E_j+u+1}b^{(j)} \right\|^2 = \tau^2\delta^2$
    - Compute $r_u^{(j)} = b^{(j)} - A\widetilde{V}_{E_j+u}\hat{y}$
  - **end**
  - $x^{(j)} = \widetilde{V}_{E_j+u}\hat{y}$
  - $E_{j+1} = E_j + u$
- **end**

end

basis vectors from the Golub-Kahan process (26) need to be stored for reorthogonalizing against appended residual vectors when solving subsequent problems in a sequence. When considering well-posed problems, the algorithms are terminated when the norm of the residual vector satisfies a tolerance criterion, while for linear discrete ill-posed problems the algorithms are terminated when the discrepancy principle (DP) is satisfied.

In our experiments with linear discrete ill-posed problems, we assume knowledge of an upper bound for the error that contaminates the data vectors $b^{(i)}$; cf. (20) and (21). The DP prescribes that an iterative method for the approximate solution of each one of the problems (1) should be terminated as soon as an iterate satisfies (22). In our experiments we let $\tau = 1.01$ in (22) and track the relative residual norms

$$\mathrm{RRN}\left(x^{(q,i)}\right) = \frac{\left\|Ax^{(q,i)} - b^{(i)}\right\|}{\|b_{true}\|},$$

where $x^{(q,i)}$ denotes the solution determined by the appropriate algorithm at iteration q when solving problem i of (1), to determine when the DP is satisfied. To evaluate the quality of the computed solutions, we compute the relative reconstructive error (RRE) defined by

$$\mathrm{RRE}\left(x^{(q,i)}\right) = \frac{\left\|x^{(q,i)} - x_{true}\right\|}{\|x_{true}\|}.$$

We measure the computational effort required by the methods in our examples by the number of iterations required. This measures the number of matrix-vector products with A (and with A^T, depending on the method).

When solving problems (1) with an ill-conditioned matrix A, we apply Tikhonov regularization both with $L = I$ and with $L \neq I$. For problems in one space dimension, we use the regularization matrix

$$L = L_1 = \begin{bmatrix} 2 & -1 & & & \\ -1 & 2 & -1 & & \\ & \ddots & \ddots & \ddots & \\ & & -1 & 2 & -1 \\ & & & -1 & 2 \end{bmatrix},$$

which is the Laplacian matrix in one space-dimension. For examples in two space-dimensions, we use the Laplacian matrix defined by

$$L = L_1 \otimes I + I \otimes L_1, \tag{38}$$

where I is the identity matrix and $\otimes$ denotes the Kronecker product. The size of L is such that left multiplication in Algorithms 3 and 5 is compatible; see [15] for details.

We first consider the application of the Arnoldi method to problems (1) with a square well-conditioned matrix from the MATLAB gallery and to problems with a very ill-conditioned matrix obtained by a modification of an example from the Regularization Tools toolbox [13]. We then turn to the testing of the FGK and RFGK methods, where for the former we consider a parallel beam tomography example and for the latter a stream of blurred images from a video.

Parter The first problem we consider is obtained by using the MATLAB gallery matrix `parter`. We set $A \in \mathbb{R}^{4000\times 4000}$. This matrix is non-symmetric and very well-conditioned; the conditioning number of A is $\kappa_2(A) \approx 4.8$. We pick 30 right-hand sides obtained by

$$b^{(i)} = Ax^{(i)},$$

where $x^{(i)}$ is a uniform sampling of the function $f^{(i)}(t) = \sin(\sigma^{(i)}t)$ in $[0, 2\pi]$ and $\sigma^{(i)} = 1 + i/30$. We compare the performance of the FA method with that of GMRES applied independently to each right-hand side, and measure the RRE and the number of iterations carried out for each right-hand side. Table 1 reports RRE values and number of iterations for a few right-hand side vectors $b^{(i)}$; the ones that are not reported are very similar to the ones reported and therefore ommitted. Figure 1(a) shows the number of iterations carried out for each right-hand side. We can observe that the GMRES method gives more accurate approximate solutions with the chosen stopping criterion than the FA method, but the latter also produces very accurate approximate solutions and is much faster. Figure 1(a) shows the number of iterations required for the FA method to be quite small after the first two problems (1) have been solved.

Shaw The matrix of this example is severely ill-conditioned. We construct the problems as follows. Discretize the integral equation discussed by Shaw [33] by a trapezoidal rule. This yields a non-symmetric matrix $A_1 \in \mathbb{R}^{64\times 64}$. MATLAB code for generating the matrix A_1 is available at [23]. We then define $A = A_1 \otimes A_1 \in \mathbb{R}^{4096\times 4096}$. Note that to evaluate matrix-vector products with A, this matrix does not have to be explicitly formed. We construct 30 exact solutions $x^{(i)}$ as the vectorized

Table 1 Parter example: RRE for FA and GMRES method at termination

Data vector i	FA (It.)	GMRES (It.)
1	$6.4991 \cdot 10^{-12}$ (77)*	$6.4991 \cdot 10^{-12}$(77)*
4	$9.0854 \cdot 10^{-6}$ (19)	$5.9404 \cdot 10^{-12}$ (77)*
7	$7.0824 \cdot 10^{-6}$ (8)	$5.4526 \cdot 10^{-12}$ (79)*
10	$4.4588 \cdot 10^{-6}$ (1)	$6.1025 \cdot 10^{-12}$ (79)*
13	$2.0136 \cdot 10^{-6}$ (1)	$5.0512 \cdot 10^{-12}$ (79)*
17	$5.6739 \cdot 10^{-7}$ (3)	$6.9302 \cdot 10^{-12}$ (76)*
20	$1.6612 \cdot 10^{-6}$ (1)	$5.6197 \cdot 10^{-12}$ (78)*
23	$1.8550 \cdot 10^{-6}$ (3)	$5.5773 \cdot 10^{-12}$ (79)*
26	$1.4089 \cdot 10^{-6}$ (6)	$5.4669 \cdot 10^{-12}$ (79)*
30	$3.8467 \cdot 10^{-7}$ (2)	$5.7744 \cdot 10^{-12}$ (77)*

The number of iterations for each method are given in parentheses. An asterisk, *, denotes the method with the smallest RRE for each data vector $b^{(i)}$

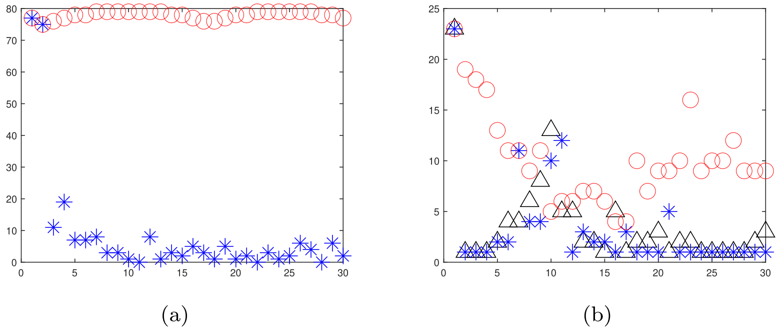

Fig. 1 Comparison of GMRES and FA, and Arnoldi-Tikhonov and RFA. Number of iterations carried out for each right-hand side. The red circles show results for the GMRES or Arnoldi-Tikhonov methods, the blue asterisks display results for the FA and RFA methods with L equal to the identity, and the black triangles report the results for RFA with L equal to the Laplacian. Panel (a) reports the results for the Parter example, while panel (b) shows results for the Shaw example

versions of uniform samplings on a 64×64 grid of the functions

$$f^{(i)}(s,t) = \exp(-\frac{\left(s/\sigma_s^{(i)}\right)^2 + \left(t/\sigma_t^{(i)}\right)^2}{\theta^{(i)}}) \left(\sin(2s) + 2\right),$$

with $(s,t) \in [-3,3]^2$, $\sigma_s^{(i)} = 1 + i/30$, $\sigma_t^{(i)} = \frac{1}{2} + i/30$, and $\theta^{(i)} = \frac{1}{2} + 2i/30$. The right-hand data vectors $b^{(i)}$ are defined by

$$b^{(i)} = Ax^{(i)} + \eta^{(i)},$$

where $\eta^{(i)}$ is a vector that simulates the presence of noise in the data. Each entry of $\eta^{(i)}$ is a realization of a Gaussian random variable with zero mean and fixed variance. We scale the vectors $\eta^{(i)}$ such that

$$\left\|\eta^{(i)}\right\| = 0.03 \left\|b^{(i)}\right\|,$$

i.e., each vector $b^{(i)}$ is contaminated by 3% noise.

We solve the problems (1) by the regularized flexible Arnoldi method of Section 3 and by solving each problem independently by the Arnoldi-Tikhonov method. The latter is equivalent to applying lines 1-10 of Algorithm 3 to solve each one of the k problems.

We let the regularization matrix L be the identity or be given by (38) and compare the RFA method with the Arnoldi-Tikhonov applied separately to each vector $b^{(i)}$. Table 2 and Fig. 1(b) report results obtained for these solution methods. Similarly as above, Table 2 reports the RREs obtained for a few vectors representative vectors $b^{(i)}$. We can observe that the RRE is smaller for the RFA method. The Laplacian regularization matrix gives reconstructions of higher quality than the identity. This is particularly true for the last slices. The computational cost is significantly smaller for

Table 2 Shaw example: RRE for the RFA and Arnoldi-Tikhonov methods at termination

Slice no.	RFA L =Identity (It.)	RFA L =Lapl. (It.)	Arnoldi-Tik. (It.)
1	1.6672 (23)*	1.6678 (23)	1.6678 (23)
4	0.2600 (1)	0.2438 (1)*	1.2756 (17)
7	0.1322 (4)	0.1047 (11)*	0.1418 (11)
10	0.0687 (13)*	0.1113 (10)	0.228 (5)
13	0.0509 (2)*	0.0862 (3)	0.0653 (7)
17	0.0830 (1)	0.0677 (3)*	0.1625 (4)
20	0.1182 (3)	0.0835 (1)*	0.1381 (9)
23	0.1584 (2)	0.1176 (1)*	0.1400 (16)
26	0.1999 (1)	0.1577 (1)*	0.2123 (10)
30	0.2520 (3)	0.2072 (1)*	0.3032 (9)

The number of iterations for each method are given in parentheses. An asterisk, *, denotes the method with the smallest RRE for each slice

the RFA method than for the Arnoldi-Tikhonov method. This can be surmised from the number of matrix-vector product evaluations required. Similarly as in the previous example, the number of iterations is small for the larger indices i.

MRI Images This example considers the performance of the FGK method described by Algorithm 4 when applied to a sequence of linear parallel beam tomography problems. We compare to the truncated GK method, which is equivalent to LSQR [25] and may be summarized by lines 1-9 in Algorithm 4. In the underlying model, incident parallel X-ray beams penetrate an object and the damping is recorded. The forward operation

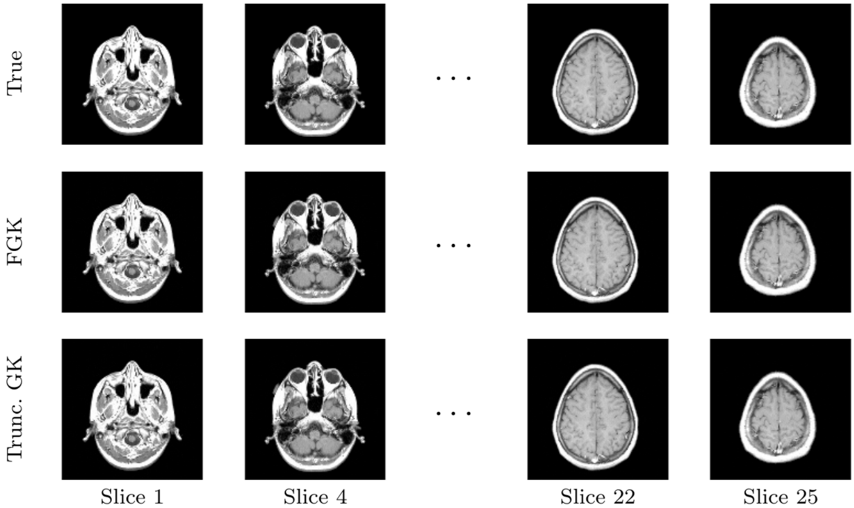

Fig. 2 MRI image example: Reconstructions of slices 1, 4, 22, and 25 determined when the discrepancy principle is satisfied for FGK and truncated GK methods and the true slices

Table 3 MRI images example: RRE at termination for the FGK and truncated GK methods at termination

Image no.	FGK (It.)	Trunc. GK (It.)
1	0.0842 (16)*	0.0842 (16)*
4	0.0860 (14)	0.0859 (17)*
7	0.0758 (10)	0.0740 (16)*
10	0.0718 (9)	0.0713 (15)*
13	0.0796 (9)*	0.0818 (15)
16	0.0814 (8)*	0.0815 (15)
19	0.0710 (8)	0.0708 (15)*
22	0.0605 (7)*	0.0610 (14)
25	0.0507(7)*	0.0507 (14)*

The number of iterations for each method are given in parentheses. An asterisk, *, denotes the method with the smallest RRE for each image

is the application of the tomography operator to a sequence of images from the MRI data set available through MATLAB's Image Processing Toolbox. This produces a sequence of sinograms that are stored in the vectors $b^{(i)}$, $i = 1, 2, \ldots, k$.

We aim to recover the spatially varying attenuation coefficients of the object of interest; see [5] for further details on the mathematical model. Using the vectorized traverse image (i.e. top-down view) of size 128×128 from the MRI data set and the default settings of the 2D X-ray tomography operator from the software package AIR Tools II, the forward operator is represented by a matrix $A \in \mathbb{R}^{32580\times 16384}$ that is stored in sparse format, where $16384 = 128^2$. Further details on the forward problem can be found in the [14]. The condition number of A is $\mathcal{O}(10^3)$ [11]. Thus this test problem is fairly well-conditioned and regularization is not necessary. We consider

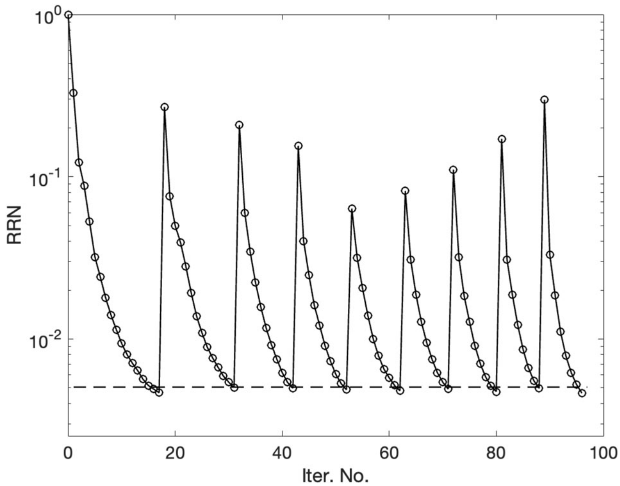

Fig. 3 MRI images example: RRN versus iteration number for the FGK method. The black dashed horizontal line represents the termination level for each sequential subproblem according to the discrepancy principle

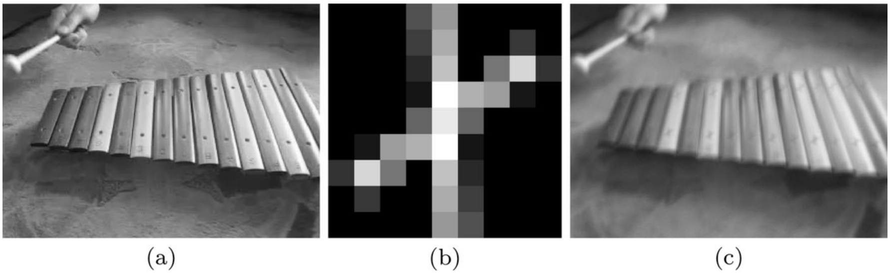

Fig. 4 Video restoration example: (a) True frame 1 image (222 × 302), (b) PSF (9 × 9), (c) blurred and 1% noised frame 1 (222 × 302)

noise contamination of 0.5% of the data vectors. The noise is represented by a vector $e \in \mathbb{R}^{32580}$ with normally distributed random entries with zero mean. The iterations for each data vector are terminated with aid of the discrepancy principle.

We consider the recovery of a sequence of 9 CT images, which we will refer to as *slices*, from the MRI data set with a gap of 3 slices to simulate the situation where the right-hand sides may not be close in norm. A subset of the true slices and their reconstructions are shown in Fig. 2. Tabulated results for both methods are displayed in Table 3. The number of iterations for FGK was smaller than when applying truncated GK to each subproblem independently. Figure 3 shows the RRN. Each spike corresponds to a new image. After the seed problem with the zero vector as the initial approximate solution is solved, the relative residual norms for subsequent images stay below 1. This illustrates the usefulness of not determining the images in the sequence independently.

Contaminated video restoration Our final example considers the restoration of a sequence of blurred and noise-contaminated video frames from the `xylophone.mp4` video available through MATLAB's Computer Vision System Toolbox. Our task is to recover a sequence of deblurred frames. Image deblurring problems are well known to give matrices that are severely ill-conditioned. This suggests that flexible regularized methods such as the ones outlined in Sections 3 and 5 should be used. We investigate the application of the RFGK method as defined in Algorithm 5 with regularization matrices given by the identity, i.e. $L = I$, as well as by the 2D Laplacian defined (38). We compare this approach to the application of the GK-Tikhonov method to each subproblem independently. This is a Tikhonov regularized method based on the truncated GK method. It is defined by applying lines 1-10 of Algorithm 5. A related method was first described in [16]. We terminate all methods with the DP for each frame.

We construct a sequence of blurred images as follows. Start with an image of $n_1 \times n_2$ pixels and blur it by using a point spread function (PSF) of size $m_1 \times m_2$ with $m_j < n_j$, $j = 1, 2$. We imposed reflective boundary conditions. Next, to simulate a real situation we cut out the boundary from the blurred image of half the size of the PSF, i.e., of size $\lceil \frac{m_j}{2} \rceil$. Noise is then added to each frame to obtain a sequence of problems of the form by (1). For a detailed description of image deconvolution; see [15].

For our example, we blur a sequence of the first 7 consecutive frames of the `\xylophone.mp4`\video transformed into grayscale with a motion blur PSF of size 9×9 pixels; see Fig. 4(b). We add 1% Gaussian noise. This particular blurring matrix has a condition number larger than 10^5. The true first frame is shown in Fig. 4(a) and (c) displays the associated blurred and noisy image. We depict a subset of the true frames and their reconstructions in Fig. 5.

RRE values for each frame for the four methods considered are recorded in Table 4. RFGK and GK-Tikhonov carry out the same number of iterations for the first frame. For all seven slices considered, The RFGK-I method, i.e., the RFGK method with $L = I$ performed the best in terms of smallest RRE values attained, and required the smallest total number of iterations to satisfy the DP. Given the relatively smooth

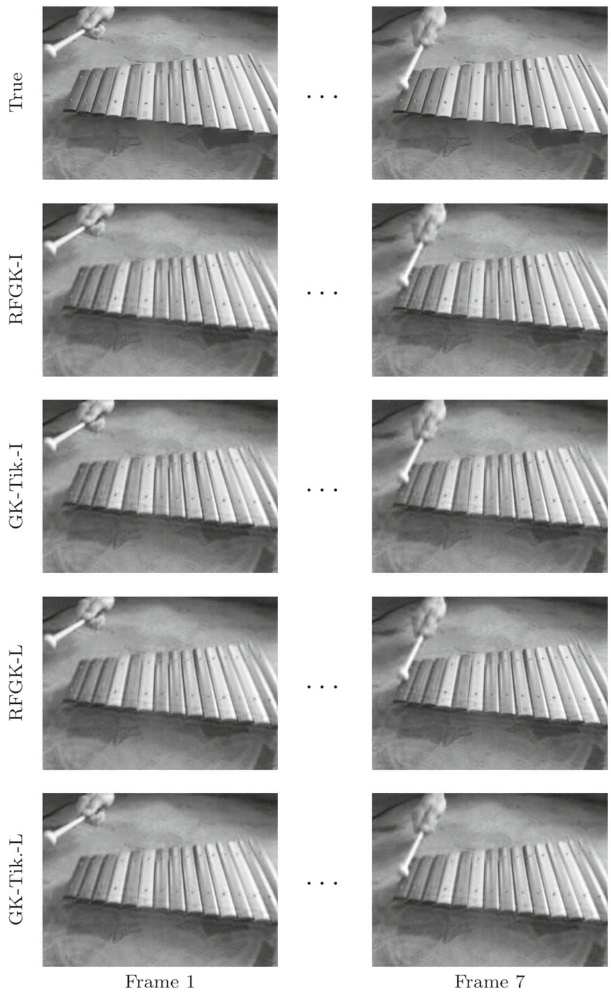

Fig. 5 Contaminated video restoration: Reconstructions of frames 1 and 7 obtained when the discrepancy principle is satisfied for the three methods considered as well as the true frames

Table 4 Contaminated video restoration: RRE for each method at termination with the discrepancy principle

Slice no.	RFGK-I (It.)	GK-Tik.-I (It.)	RFGK-L (It.)	GK-Tik.-L (It.)
1	0.0571 (7)*	0.0571 (7)*	0.0576 (7)	0.0576 (7)
2	0.0531 (5)*	0.0574 (7)	0.0537 (5)	0.0579 (7)
3	0.0511 (3)*	0.0572 (7)	0.0513 (3)	0.0577 (7)
4	0.0507 (4)*	0.0559 (7)	0.0510 (4)	0.0564 (7)
5	0.0483 (4)*	0.0556 (7)	0.0488 (4)	0.0562 (7)
6	0.0494 (4)*	0.0570 (7)	0.0496 (4)	0.0576 (7)
7	0.0487 (3)*	0.0565 (7)	0.0488 (3)	0.0570 (7)

The number of iterations for each method are given in parentheses. An asterisk, *, denotes the method with the smallest RRE for each slice

solution of this particular example compared to the solution of the *Shaw* problem, we expected that RFGK-L method, i.e., the RFGK method with L the 2D Laplacian, would not perform as well as RFGK-I. This is confirmed by the RRE values for each frame obtained with the RFGK-I and RFGK-L methods. Selected reconstructions for each method are shown in Fig. 5.

7 Conclusion

We have described several methods based on the flexible Arnoldi decomposition or the flexible Golub-Kahan decomposition to solve sequences of linear systems of equations or linear least-squares problems with the same matrix. When the matrices involved in the linear systems are very ill-conditioned, we employed regularization. Computed examples show the proposed methods to outperform their non-flexible counterparts in terms of both accuracy and computational cost.

Acknowledgements The authors would like to thank Enyinda Onunwor for some initial discussions that lead to this paper. A.B. and M.D. are part of the GNCS group of IN- dAM and A.B. is partially supported by INdAM-GNCS 2024 Project "Algebra lineare numerica per problemi di grandi dimensioni: aspetti teorici e applicazioni" (CUP E53C23001670001). A.B. is partially supported by the PRIN 2022 PNRR project no. P2022PMEN2 financed by the European Union - NextGenerationEU and by the Italian Ministry of University and Research (MUR). A.B. is partially supported by the PRIN 2022 project "Inverse Problems in the Imaging Sciences (IPIS)" (2022ANC8HL) financed by the European Union - NextGenerationEU and by the Italian Ministry of University and Research (MUR). A.B.'s work is partially founded by Fondazione di Sardegna, Progetto biennale bando 2021, "Computa- tional Methods and Networks in Civil Engineering (COMANCHE)". L.O. is par- tially supported by the U.S. National Science Foundation, Grants DMS-2038118 and DMS-2208294.

Author Contributions All authors contributed equally to the paper.

Funding See acknowledgement.

Data Availability No datasets were generated or analysed during the current study.

Declarations

Competing interests The authors declare no competing interests.

References

1. Abdel-Rehim, A.M., Morgan, R.B., Wilcox, W.: Improved seed methods for symmetric positive definite linear equations with multiple right-hand sides. Numer. Linear Algebra Appl. **21**, 453–471 (2014)
2. Al Daas, H., Grigori, L., Hénon, P., Ricoux, P.: Recycling Krylov subspaces and reducing deflation subspaces for solving sequence of linear systems. ACM Trans. Math. Software **47**, 1–30 (2021)
3. Bentbib, A.H., El Guide, M., Jbilou, K., Onunwor, E., Reichel, L.: Solution methods for linear discrete ill-posed problems for color image restoration. BIT Numer. Math. **58**, 555–578 (2018)
4. Buccini, A., Reichel, L.: Software for limited memory restarted ℓ^p-ℓ^q minimization methods using generalized Krylov subspaces. Electron. Trans. Numer. Anal. **61**, 66–91 (2024)
5. Buzug, T.M.: Computed Tomography. Springer, Berlin (2008)
6. Calvetti, D., Reichel, L.: Application of a block modified Chebyshev algorithm to the iterative solution of symmetric linear systems with multiple right hand side vectors. Numer. Math. **68**, 3–16 (1994)
7. Chan, T.F., Wan, W.L.: Analysis of projection methods for solving linear systems with multiple right-hand sides. SIAM J. Sci. Comput. **18**, 1698–1721 (1997)
8. Chung, J., Gazzola, S.: Flexible Krylov methods for ℓ_p regularization. SIAM J. Sci. Comput. **41**, S149–S171 (2019)
9. Dykes, L., Reichel, L.: Simplified GSVD computations for the solution of linear discrete ill-posed problems. J. Comput. Appl. Math. **255**, 15–27 (2013)
10. Dykes, L., Noschese, S., Reichel, L.: Rescaling the GSVD with application to ill-posed problems. Numer. Algorithms **68**, 531–545 (2015)
11. Gazzola, S., Hansen, P.C., Nagy, J.G., Tools, I.R.: A MATLAB package of iterative regularization methods and large-scale test problems. Numer. Algorithms **81**, 773–811 (2019)
12. Hansen, P.C.: Rank-Deficient and Discrete Ill-Posed Problems. SIAM, Philadelphia (1998)
13. Hansen, P.C.: Regularization tools version 4.0 for Matlab 7.3. Numer. Algorithms **46**(5), 189–194 (2007)
14. Hansen, P.C., Jørgensen, J.S., Tools, A.I.R., II.: Algebraic iterative reconstruction methods, improved implementation. Numer. Algorithms **79**, 107–137 (2018)
15. Hansen, P.C., Nagy, J.G., O'Leary, D.P.: Deblurring Images: Matrices, Spectra, and Filtering. SIAM, Philadelphia (2006)
16. Hochstenbach, M.E., Reichel, L.: An iterative method for Tikhonov regularization with a general linear regularization operator. J. Integral Equations Appl. **22**, 463–480 (2010)
17. Huang, G., Reichel, L., Yin, F.: On the choice of subspace for large-scale Tikhonov regularization problems in general form. Numer. Algorithms **81**, 33–55 (2019)
18. Jiang, J., Chung, J., de Sturler, E.: Hybrid projection methods with recycling for inverse problems. SIAM J. Sci. Comp. **43**, S146–S172 (2021)
19. Kindermann, S.: Convergence analysis of minimization-based noise level-free parameter choice rules for linear ill-posed problems. Electron. Trans. Numer. Anal. **38**, 233–257 (2011)
20. Kindermann, S., Raik, K.: A simplified L-curve method as error estimator. Electron. Trans. Numer. Anal. **53**, 217–238 (2020)
21. Lampe, J., Reichel, L., Voss, H.: Large-scale Tikhonov regularization via reduction by orthogonal projection. Linear Algebra Appl. **436**, 2845–2865 (2012)
22. Lanza, A., Morigi, S., Reichel, L., Sgallari, F.: A generalized Krylov subspace method for ℓ_p-ℓ_q minimization. SIAM J. Sci. Comput. **37**, S30–S50 (2015)
23. Neuman, A., Reichel, L., Sadok, H.: Algorithms for range restricted iterative methods for linear discrete ill-posed problems. Numer. Algorithms **59**, 325–331 (2012)
24. Onisk, L., Reichel, L., Sadok, H.: Numerical considerations of block GMRES methods when applied to linear discrete ill-posed problems. J. Comput. Appl. Math. **430**, 115262 (2023)
25. Paige, C.C., Saunders, M.A.: LSQR: An algorithm for sparse linear equations and sparse least squares. ACM Trans. Math. Softw. **8**, 43–71 (1982)
26. Parlett, B.N.: A new look at the Lanczos algorithm for solving symmetric systems of linear equations. Linear Algebra Appl. **29**, 323–346 (1980)

27. Pasha, M., Saibaba, A.K., Gazzola, S., Español, M.I., de Sturler, E.: A computational framework for edge-preserving regularization in dynamic inverse problems. Electron. Trans. Numer. Anal. **58**, 486–516 (2023)
28. Reichel, L., Rodriguez, G.: Old and new parameter choice rules for discrete ill-posed problems. Numer. Algorithms **63**, 65–87 (2013)
29. Reichel, L., Ye, Q.: Breakdown-free GMRES for singular systems. SIAM J. Matrix Anal. Appl. **26**, 1001–1021 (2005)
30. Saad, Y.: On the Lanczos method for solving symmetric linear systems with several right-hand sides. Math. Comput. **48**, 651–662 (1987)
31. Saad, Y.: A flexible inner-outer preconditioned GMRES algorithm. SIAM J. Sci. Comput. **14**, 461–469 (1993)
32. Saad, Y.: Iterative Methods for Sparse Linear Systems, 2nd edn. SIAM, Philadelphia (2003)
33. Shaw, C.B., Jr.: Improvements of the resolution of an instrument by numerical solution of an integral equation. J. Math. Anal. Appl. **37**, 83–112 (1972)
34. Simoncini, V., Gallopoulos, E.: A hybrid block GMRES method for nonsymmetric systems with multiple right-hand sides. J. Comput. Appl. Math. **66**, 457–469 (1997)
35. Soodhalter, K.M., de Sturler, E., Kilmer, M.E.: A survey of subspace recycling iterative methods. GAMM-Mitt. **43**, e202000016 (2020)

Authors and Affiliations

Alessandro Buccini[1] · Marco Donatelli[2] · Lucas Onisk[3] · Lothar Reichel[4]

✉ Lucas Onisk
lonisk@emory.edu

Alessandro Buccini
alessandro.buccini@unica.it

Marco Donatelli
marco.donatelli@uninsubria.it

Lothar Reichel
reichel@math.kent.edu

1 Dipartimento di Matematica e Informatica, Università di Cagliari, 09124 Cagliari, Italy

2 Dipartimento di Scienza e Alta Tecnologia, Università dell'Insubria, 22100 Como, Italy

3 Department of Mathematics, Emory University, Atlanta, GA 30322, USA

4 Department of Mathematical Sciences, Kent State University, Kent, OH 44242, USA