Numerical Algorithms (2025) 100:1617–1637
https://doi.org/10.1007/s11075-025-02072-2

RESEARCH

Projected iterated Tikhonov in general form with adaptive choice of the regularization parameter

Alessandro Buccini[1] · Silvia Gazzola[2] · Lucas Onisk[3] · Mirjeta Pasha[4] · Lothar Reichel[5]

Received: 9 January 2025 / Accepted: 12 April 2025 / Published online: 6 May 2025

Abstract
Tikhonov regularization is commonly used in the solution of linear discrete ill-posed problems. It is known that iterated Tikhonov regularization often produces approximate solutions of higher quality than (standard) Tikhonov regularization. This paper discusses iterated Tikhonov regularization for large-scale problems with a general regularization matrix. Specifically, the original problem is reduced to small size by application of a fairly small number of steps of the Arnoldi or Golub-Kahan processes, and iterated Tikhonov is applied to the reduced problem. The regularization parameter is determined by using an extension of a technique first described by Donatelli and Hanke for quite special coefficient matrices. Convergence of the method is established and computed examples illustrate its performance.

Keywords Inverse problems · Tikhonov regularization · Iterative regularization

1 Introduction

We consider the solution of minimization problems of the form

$$\min_{\mathbf{x}\in\mathbb{R}^n} \|\mathbf{A}\mathbf{x}-\mathbf{b}\|_2^2, \tag{1}$$

where $\mathbf{A} \in \mathbb{R}^{m\times n}$ is a large severely ill-conditioned matrix whose singular values decrease to zero with increasing index with no significant gap, and the vector $\mathbf{b} \in \mathbb{R}^m$ represents data. Throughout this paper $\|\cdot\|_2$ denotes the Euclidean vector norm. In the minimization problems considered, the vector $\mathbf{b}$ is not available; only a noise-contaminated version $\mathbf{b}^\delta = \mathbf{b} + \mathbf{e}$ is known. The contamination $\mathbf{e} \in \mathbb{R}^m$ may, for instance, stem from measurement inaccuracies. Thus, we are faced with the solution

Dedicated to our friends Michela Redivo-Zaglia and Hassane Sadok.

All authors contributed equally to this work

Extended author information available on the last page of the article of the least squares problem

$$\min_{\mathbf{x}\in\mathbb{R}^n} \left\|\mathbf{A}\mathbf{x}-\mathbf{b}^\delta\right\|_2^2, \tag{2}$$

and would like the computed solution to be an accurate approximation of the solution of minimal Euclidean norm of (1). We denote the latter solution by $\mathbf{x}^\dagger$. However, due to the severe ill-conditioning of $\mathbf{A}$ and the presence of noise in $\mathbf{b}^\delta$, the least-squares solution of minimal Euclidean norm of (2) typically is a poor approximation of $\mathbf{x}^\dagger$.

In order to be able to determine a useful approximation of $\mathbf{x}^\dagger$, the least-squares problem (2) is replaced by a nearby problem obtained by Tikhonov regularization, whose solution is less sensitive to the error in $\mathbf{b}^\delta$ than the solution of (2). Tikhonov regularization replaces the minimization problem (2) by the problem

$$\min_{\mathbf{x}\in\mathbb{R}^n} \left\|\mathbf{A}\mathbf{x}-\mathbf{b}^\delta\right\|_2^2 + \mu\left\|\mathbf{L}\mathbf{x}\right\|_2^2, \tag{3}$$

where $\mu > 0$ is a regularization parameter and $\mathbf{L} \in \mathbb{R}^{s\times n}$ is a regularization matrix. We will assume that $1 \le s \le \min\{m, n\}$, but it is easy to remove the upper bound for s. The null spaces $\mathcal{N}(\mathbf{A})$ and $\mathcal{N}(\mathbf{L})$ of the matrices $\mathbf{A}$ and $\mathbf{L}$, respectively, are assumed to satisfy

$$\mathcal{N}(\mathbf{A}) \cap \mathcal{N}(\mathbf{L}) = \{\mathbf{0}\}, \tag{4}$$

so that problem (3) has a unique solution for any $\mu > 0$. For a recent review of many solution methods for linear discrete ill-posed problems, including Tikhonov regularization, we refer to [1].

It is well-known (see, e.g., [2]) that the solution of (3) may be over-smoothed, i.e., the details of $\mathbf{x}^\dagger$ may not show in the solution of (3). A possible approach to reduce over-smoothing is to apply an iterated Tikhonov regularization method. Iterated Tikhonov regularization may be considered a preconditioned version of the iterative Landweber solution method or as a refinement technique of (non-iterated) Tikhonov regularization. In either case, the resulting iterates are of the form

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \left(\mathbf{A}^T\mathbf{A} + \mu_k\mathbf{L}^T\mathbf{L}\right)^{-1}\mathbf{A}^T\left(\mathbf{b}^\delta - \mathbf{A}\mathbf{x}^{(k)}\right), \quad k = 0, 1, \dots, \tag{5}$$

where $\mathbf{x}^{(0)}$ is an initial approximation of $\mathbf{x}^\dagger$. When $\mathbf{x}^{(0)} = \mathbf{0}$ and $k = 0$, eq. (5) simplifies to (3). Discussions on the iterated Tikhonov method and illustrations of its performance can be found in, e.g., [2–9]. These references discuss the choices of the regularization parameters $\mu_k > 0$, of the regularization matrix $\mathbf{L}$, and of the subspaces in which approximate solutions of (5) are determined. The need to choose solution subspaces arises when solving large-scale problems to avoid the solution of a linear system of equations with a large matrix in every iteration (5).

The convergence of the iterates $\mathbf{x}^{(k+1)}$ determined by (5) for increasing values of k depends on the sequence of regularization parameters $\{\mu_k\}_{k\ge 0}$. If the parameters $\mu_k > 0$ satisfy

$$\sum_{k=0}^{\infty} \mu_k^{-1} = \infty,$$

then the iterates converge to a solution of (2) that is closest to $\mathbf{x}^{(0)}$ in some norm; see [4] and references therein. For some choices of the parameters μ_k, the iterated Tikhonov method (5) is "semiconvergent", i.e., the iterates $\mathbf{x}^{(k+1)}$, $k = 0, 1, \ldots$, determined by the method approach the desired solution $\mathbf{x}^\dagger$ when the index k increases and is sufficiently small, but for larger index values the iterates diverge from $\mathbf{x}^\dagger$. To avoid "divergence", the iterations are terminated sufficiently early.

If a fairly accurate estimate for the bound $\delta > 0$ for the noise in $\mathbf{b}^\delta$ is available, i.e.,

$$\left\|\mathbf{b}^\delta - \mathbf{b}\right\|_2 \leq \delta, \tag{6}$$

then one can use the discrepancy principle (DP) to determine the regularization parameters μ_k and to decide when to terminate the iterations. A way to apply the DP for iterated Tikhonov regularization is described below. Under the assumption that $\mathbf{b} \in \mathrm{range}(\mathbf{A})$, the DP prescribes that the iterations be terminated at iteration k_{DP}, where

$$k_{\mathrm{DP}} = \min\left\{k : \left\|\mathbf{A}\mathbf{x}^{(k)} - \mathbf{b}^\delta\right\|_2 \leq \tau\delta\right\} \tag{7}$$

and $\tau > 1$ is a user-specified constant (typically chosen close to 1). This stopping criterion is quite natural since $\left\|\mathbf{A}\mathbf{x}^\dagger - \mathbf{b}\right\|_2 \leq \tau\delta$. If a fairly accurate estimate for δ is not available, then other stopping criteria can be used, such as criteria that are based on the L-curve or generalized cross-validation; see, e.g., [1, 10–12] for discussions.

The main difficulties when applying the iterations (5) are that: i) an imprudent choice of the parameters μ_k may produce slow or no convergence of the iterates $\mathbf{x}^{(k)}$ to $\mathbf{x}^\dagger$ as k increases, and ii) a large linear system of equations has to be solved in each iteration.

It is the aim of this paper to present a convergence analysis for the iterated Tikhonov method based on an extension of the technique employed by Donatelli and Hanke [13] in their analysis of the Approximated Iterated Tikhonov (AIT) method with $\mathbf{L}$ replacing the identity matrix $\mathbf{I}$. The AIT method replaces the matrix $\mathbf{A}$ with a spectrally equivalent matrix $\mathbf{C}$ in the preconditioner $\left(\mathbf{A}^T\mathbf{A} + \mu_k\mathbf{I}\right)^{-1}$ (appearing in (5), with $\mathbf{L} = \mathbf{I}$). A way to determine suitable regularization parameters μ_k is also described. Subsequently, the AIT algorithm was extended to special regularization matrix pairs $\{\mathbf{A}, \mathbf{L}\}$ where $\mathbf{L}$ is not required to be the identity. The convergence result in [13] was sharpened in [5]; see also [14]. To allow for more general matrices $\mathbf{A}$ than those discussed in [13], a new version of the AIT method, referred to as the Iterated Arnoldi Tikhonov (IAT) method, was proposed in [6]. This method projects the problem to be solved into a Krylov subspace of fairly small dimension. However, since the Arnoldi procedure was selected to generate such Krylov subspace, the method only can handle square matrices $\mathbf{A}$ and requires $\mathbf{L} = \mathbf{I}$.

This paper proposes new variations of the AIT method that allow efficient solution of (5) with $\mathbf{L} \neq \mathbf{I}$ by projections onto Krylov subspaces generated by either the Arnoldi or the Golub-Kahan bidiagonalization algorithms. We refer to such solvers as Projected Iterated Tikhonov (PIT) methods. They can be regarded as extensions of the IAT method in [6]. In particular, to apply the methods proposed in this paper, the matrix $\mathbf{A}$ is not required to be square. The new methods differ from those in [6, 8, 9] in that they allow multiple iterations of a projected problem whose dimension is determined by applying the discrepancy principle (7). We experimentally illustrate that allowing a suitable number of iterated Tikhonov iterations of the projected problem may improve the quality of the computed solution.

This paper is organized as follows. Section 2 reviews some projection methods that form the basis for the new PIT method and describes the latter. Section 3 presents the convergence analysis for PIT. Computed results are presented in Section 4, alongside some algorithmic details about PIT. Concluding remarks can be found in Section 5.

2 Projection methods and the new PIT method

Let the columns $\mathbf{V}_p \in \mathbb{R}^{n\times p}$ form an orthonormal basis for the subspace $\mathcal{V}_p$ in which we would like to determine an approximate solution of problem (2). Typically, $1 \leq p \ll \min\{m, n\}$. The methods we use to determine the columns of $\mathbf{V}_p$ simultaneously compute a decomposition of $\mathbf{A}\mathbf{V}_p \in \mathbb{R}^{m\times p}$, which involves a small matrix of order $(p+1)\times p$ that can be interpreted as a projection of $\mathbf{A}$ into the subspace $\mathcal{V}_p$. We consider the following methods:

(i) Let $\mathbf{A} \in \mathbb{R}^{n\times n}$. Application of $p \leq n$ steps of the Arnoldi process to $\mathbf{A}$ with initial vector $\mathbf{b}^\delta$ generically gives the Arnoldi decomposition

$$\mathbf{A}\mathbf{V}_p = \mathbf{V}_{p+1}\mathbf{H}_{p+1,p}, \tag{8}$$

where $\mathbf{H}_{p+1,p} \in \mathbb{R}^{(p+1)\times p}$ is an upper Hessenberg matrix with positive subdiagonal entries. The columns of $\mathbf{V}_p$ span the Krylov subspace

$$\mathcal{K}_p(\mathbf{A}, \mathbf{b}^\delta) = \mathrm{span}\{\mathbf{b}^\delta, \mathbf{A}\mathbf{b}^\delta, \ldots, \mathbf{A}^{p-1}\mathbf{b}^\delta\},$$

with the first column proportional to $\mathbf{b}^\delta$. We note that if $\mathbf{A}$ is symmetric, then the Arnoldi process reduces to the symmetric Lanczos process and $\mathbf{H}_{p+1,p}$ becomes a symmetric tridiagonal matrix.

(ii) Let $\mathbf{A} \in \mathbb{R}^{m\times n}$ be a possibly rectangular matrix. We allow $m \geq n$ as well as $m < n$. Application of $p \leq \min\{m, n\}$ steps of the Golub-Kahan process to $\mathbf{A}$ with initial vector $\mathbf{b}^\delta$ generically determines the decomposition

$$\mathbf{A}\mathbf{V}_p = \mathbf{U}_{p+1}\mathbf{B}_{p+1,p}, \tag{9}$$

as well as an analogous decomposition that involves $\mathbf{A}^T$. Here $\mathbf{B}_{p+1,p} \in \mathbb{R}^{(p+1)\times p}$ is a lower bidiagonal matrix, and the matrices $\mathbf{U}_{p+1} \in \mathbb{R}^{m\times(p+1)}$ and $\mathbf{V}_p \in \mathbb{R}^{n\times p}$ have orthonormal columns. The columns of $\mathbf{U}_{p+1}$ span the Krylov subspace $\mathcal{K}_{p+1}(\mathbf{A}\mathbf{A}^T, \mathbf{b}^\delta)$ with the first column proportional to $\mathbf{b}^\delta$; the columns of $\mathbf{V}_p$ span the Krylov subspace $\mathcal{K}_p(\mathbf{A}^T\mathbf{A}, \mathbf{A}^T\mathbf{b}^\delta)$. Here and throughout this paper the superscript ^{T} denotes transposition.

The remainder of this section describes our solution method for (5) when the matrix $\mathbf{A}$ is reduced by the Golub-Kahan process (9). The modifications required when $\mathbf{A}$

instead is reduced by the Arnoldi process (8) are straightforward. Projecting the problem (2) into the low-dimensional solution subspace $\mathcal{V}_p$ generated by applying p steps of the Golub-Kahan process to $\mathbf{A}$ with initial vector $\mathbf{b}^\delta$, cf. (9), gives the reduced problem

$$\min_{\mathbf{y}\in\mathbb{R}^p} \left\|\mathbf{B}_{p+1,p}\mathbf{y} - \widetilde{\mathbf{b}}^\delta\right\|_2^2, \tag{10}$$

where $\widetilde{\mathbf{b}}^\delta = \left\|\mathbf{b}^\delta\right\|_2 \mathbf{e}_1 \in \mathbb{R}^{p+1}$ and $\mathbf{e}_1 = [1, 0, \ldots, 0]^T$ denotes the first column of the identity matrix of order $p+1$. In the rare event when the decomposition on the right-hand side of (9) does not exist for the chosen value of p (due to breakdown of the computations), we replace p by the largest possible integer $p' \leq p$ such that

$$\mathbf{A}\mathbf{V}_{p'} = \mathbf{U}_{p'}\mathbf{B}_{p',p'},$$

where $\mathbf{B}_{p',p'} \in \mathbb{R}^{p'\times p'}$ is of upper bidiagonal form. The computations are simplified in this situation. We will not dwell on the details and henceforth assume that the decompositions on the right-hand side of (8) exist, referring to [15, 16] for analysis of breakdowns of the Golub-Kahan and Arnoldi processes.

Let

$$\mathbf{L}\mathbf{V}_p = \mathbf{Q}\mathbf{R} \tag{11}$$

be the economical QR factorization of $\mathbf{L}\mathbf{V}_p$, where $\mathbf{Q} \in \mathbb{R}^{s\times p}$ has orthonormal columns and $\mathbf{R} \in \mathbb{R}^{p\times p}$ is upper triangular. We will assume that p is large enough so that a solution that satisfies the DP can be found in the subspace $\mathcal{V}_p$ and that the matrix $\mathbf{R}$ in (11) is nonsingular. In fact, in many applications, the matrix $\mathbf{R}$ is not very ill-conditioned.

We are in a position to describe our algorithm for the solution of (5). Since this is detailed for the Golub-Kahan bidiagonalization case, it will be referred to as Projected Iterated Tikhonov - Golub-Kahan-Bidiagonalization (PIT-GKB). Let $\mathbf{y}^{(0)} \in \mathbb{R}^p$ be an initial approximate solution and fix $\rho \in \left(0, \frac{1}{2}\right)$ and $q \in (2\rho, 1)$. Define

$$\tau = \frac{2-\rho}{2+\rho}$$

and let $\mathbf{r}^{(k)}$ denote the residual at step k, i.e.,

$$\mathbf{r}^{(k)} = \widetilde{\mathbf{b}}^\delta - \mathbf{B}_{p+1,p}\mathbf{y}^{(k)}, \quad k = 0, 1, 2, \ldots .$$

Introduce the parameters

$$\tau_k = \frac{\left\|\mathbf{r}^{(k)}\right\|_2}{\delta} \quad \text{and} \quad q_k = \max\left\{q, 2\rho + \frac{1+\rho}{\tau_k}\right\}, \tag{12}$$

and define the vector

$$\mathbf{h}^{(k)} = \left(\mathbf{B}_{p+1,p}^T\mathbf{B}_{p+1,p} + \mu_k\mathbf{R}^T\mathbf{R}\right)^{-1}\mathbf{B}_{p+1,p}^T\mathbf{r}^{(k)}, \tag{13}$$

where the regularization parameter μ_k is chosen so that the vector $\mathbf{h}^{(k)}$ satisfies

$$\left\|\mathbf{r}^{(k)} - \mathbf{B}_{p+1,p}\mathbf{h}^{(k)}\right\|_2 = q_k \left\|\mathbf{r}^{(k)}\right\|_2. \tag{14}$$

This requires the use of a zero-finder. We remark that in actual computations, we do not form the matrix $\mathbf{B}_{p+1,p}^T\mathbf{B}_{p+1,p} + \mu_k\mathbf{R}^T\mathbf{R}$ to compute $\mathbf{h}^{(k)}$; instead we evaluate the generalized singular value decomposition (GSVD) of the matrix pair $\{\mathbf{B}_{p+1,p}, \mathbf{R}\}$ and use it to compute $\mathbf{h}^{(k)}$ for specified values of the parameter μ_k; see below for details.

Having determined $\mathbf{h}^{(k)}$, we construct the new approximate projected solution

$$\mathbf{y}^{(k+1)} = \mathbf{y}^{(k)} + \mathbf{h}^{(k)}. \tag{15}$$

The iterations are terminated as soon as the $(k+1)^{st}$ residual $\mathbf{r}^{(k+1)}$ satisfies the DP, i.e., as soon as

$$\left\|\mathbf{r}^{(k+1)}\right\|_2 \le \tau\delta.$$

This determines the index $k_{\mathrm{DP}} = k+1$. The associated computed approximate solution of (3) is given by

$$\mathbf{x}^* = \mathbf{V}_p\mathbf{y}^{(k_{\mathrm{DP}})}.$$

Note that the solution subspace $\mathcal{V}_p$ is chosen independently of $\mathbf{L}$. This approach to choosing solution subspace was possibly first applied in [17]. Computed examples reported in [18] show this choice of solution subspace to perform well.

The PIT-GKB algorithm is summarized in Algorithm 1. When both PIT-GKB and PIT-A (i.e., PIT based on the Arnoldi algorithm) are applied with $\mathbf{L} = \mathbf{I}$ (5) (and, therefore, $\mathbf{R} = \mathbf{I}$ in (13)), e.g., with Tikhonov in standard form, they are referred to by the acronyms PIT-A-sf and PIT-GKB-sf, respectively. Note that PIT-A-sf coincides with the IAT method in [6].

3 Convergence analysis of PIT

In this section, we show a number of results that are helpful for analyzing the theoretical properties of the new PIT methods and show their convergence. Here, we tailor our derivations to the PIT-GKB case, but these can be easily modified to handle the PIT-A case similarly as in [6]. Introduce the (unknown) solution of the reduced problem (10) with $\widetilde{\mathbf{b}}^\delta$ replaced with $\widetilde{\mathbf{b}}$,

$$\mathbf{y}^\dagger = \mathbf{B}_{p+1,p}^\dagger\widetilde{\mathbf{b}},$$

where $\widetilde{\mathbf{b}} = \mathbf{U}_{p+1}^T\mathbf{b}$, noting that the bound δ for the norm of the noise in the data vector $\mathbf{b}^\delta$ also holds for the projected data vectors, as

$$\left\|\widetilde{\mathbf{b}}^\delta - \widetilde{\mathbf{b}}\right\|_2 = \left\|\mathbf{U}_{p+1}^T(\mathbf{b}^\delta - \mathbf{b})\right\|_2 \le \left\|\mathbf{U}_{p+1}^T\right\|_2 \left\|\mathbf{b}^\delta - \mathbf{b}\right\|_2 \le \delta.$$

Introduce the "reduced error"

$$\mathbf{e}^{(k)} = \mathbf{y}^\dagger - \mathbf{y}^{(k)}.$$

Algorithm 1 PIT-GKB Algorithm.

Input: $\mathbf{A} \in \mathbb{R}^{m\times n}$, $\mathbf{b}^{\delta} \in \mathbb{R}^{n}$, $\mathbf{L} \in \mathbb{R}^{s\times n}$, $\rho \in (0, \frac{1}{2})$, $\mathbf{y}^{(0)} \in \mathbb{R}^{p}$, $p \in \mathbb{Z}^{+}$, and $q \in (2\rho, 1)$
Output: $\mathbf{x}^{(k)} \in \mathbb{R}^{n}$
Set $\mathbf{u}_1 = \mathbf{b}^{\delta}/\|\mathbf{b}^{\delta}\|$
Compute p steps of GKB process: $\mathbf{A}\mathbf{V}_p = \mathbf{U}_{p+1}\mathbf{B}_{p+1,p}$
Compute QR factorization: $[\sim, \mathbf{R}] = \mathrm{qr}\left(\mathbf{L}\mathbf{V}_p\right)$
$\tau = \frac{1+2\rho}{1-2\rho}$
$\widetilde{\mathbf{b}}^{\delta} = \left\|\mathbf{b}^{\delta}\right\|_2 \mathbf{e}_1 \in \mathbb{R}^{p+1}$
$\mathbf{r}^{(0)} = \widetilde{\mathbf{b}}^{\delta} - \mathbf{B}_{p+1,p}\mathbf{y}^{(0)}$
for $k = 0, 1, 2, \ldots$ **do**
  $\tau_k = \|\mathbf{r}^{(k)}\|_2/\delta$
  $q_k = \max\left\{q,\ 2\rho + \frac{1+\rho}{\tau_k}\right\}$
  Determine μ_k s.t. $\left\|\mathbf{r}^{(k)} - \mathbf{B}_{p+1,p}\mathbf{h}^{(k)}\right\|_2 = q_k \left\|\mathbf{r}^{(k)}\right\|_2$
  where $\mathbf{h}^{(k)} = \left(\mathbf{B}_{p+1,p}^T\mathbf{B}_{p+1,p} + \mu_k\mathbf{R}^T\mathbf{R}\right)^{-1}\mathbf{B}_{p+1,p}^T\mathbf{r}^{(k)}$
  $\mathbf{y}^{(k+1)} = \mathbf{y}^{(k)} + \mathbf{h}^{(k)}$
  $\mathbf{r}^{(k+1)} = \widetilde{\mathbf{b}}^{\delta} - \mathbf{B}_{p+1,p}\mathbf{y}^{(k+1)}$
  if $\left\|\mathbf{r}^{(k+1)}\right\|_2 \le \tau\delta$ **then**
    $\mathbf{x}^{(k)} = \mathbf{V}_p\mathbf{y}^{(k+1)}$
  end
end

Lemma 1 *With the notation and assumptions above, it holds, for* $\left\|\mathbf{r}^{(k)}\right\|_2 \le \tau\delta$, *that*

$$\left\|\mathbf{r}^{(k)} - \mathbf{B}_{p+1,p}\mathbf{e}^{(k)}\right\|_2 \le \left(\rho + \frac{1+\rho}{\tau_k}\right)\left\|\mathbf{r}^{(k)}\right\|_2 < (1+\rho)\left\|\mathbf{r}^{(k)}\right\|_2.$$

Proof The result can be shown in the same manner as [13, Proposition 1]. We omit the details. □

We turn to the GSVD of the matrix pair $\{\mathbf{B}_{p+1,p}, \mathbf{R}\}$, defined as

$$\begin{cases} \mathbf{B}_{p+1,p} &= \mathbf{U}\boldsymbol{\Sigma}\mathbf{X}^T, \\ \quad\mathbf{R} &= \mathbf{V}\boldsymbol{\Lambda}\mathbf{X}^T, \end{cases} \tag{16}$$

where the matrices $\mathbf{U} \in \mathbb{R}^{(p+1)\times p}$ and $\mathbf{V} \in \mathbb{R}^{p\times p}$ have orthonormal columns, the matrix $\mathbf{X} \in \mathbb{R}^{p\times p}$ is non-singular, and the diagonal entries of the matrices

$$\boldsymbol{\Sigma} = \mathrm{diag}(\sigma_1, \sigma_2, \ldots, \sigma_p) \in \mathbb{R}^{p\times p}, \quad \boldsymbol{\Lambda} = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_p) \in \mathbb{R}^{p\times p},$$

are nonnegative and such that

$$\sigma_j^2 + \lambda_j^2 = 1, \quad j = 1, 2, \ldots, p;$$

see, e.g., [15] for details. Some computational aspects of the evaluation of the GSVD are discussed in [19].

Define the matrix

$$\hat{\mathbf{L}} = \mathbf{\Lambda}\mathbf{X}^T = \mathbf{V}^T\mathbf{R}$$

and note that, since $\mathbf{X}$ in (16) is non-singular and we assumed that $\mathbf{R}$ is invertible, the matrix $\hat{\mathbf{L}}$ has a trivial null space. We therefore can define the norm

$$\|\mathbf{y}\|_{\hat{\mathbf{L}}} = \left\|\hat{\mathbf{L}}\mathbf{y}\right\|_2. \tag{17}$$

This gives the following result, which guarantees monotonic decrease in the $\hat{\mathbf{L}}$ norm (defined in (17)) of the projected error vectors $\mathbf{e}^{(k)}$, $k = 0, 1, 2 \ldots$.

Proposition 2 *With the notation and assumptions above, assume that*

$$\mathbf{U}\mathbf{U}^T\mathbf{e}_1 = \mathbf{e}_1, \tag{18}$$

where $\mathbf{U}$ *is defined in* (16) *and* $\mathbf{e}_1$ *is the first column of the identity matrix of order* $p+1$. *We have that, while* $\left\|\mathbf{r}^{(k)}\right\|_2 \le \tau\delta$, *it holds*

$$\left\|\mathbf{e}^{(k)}\right\|_{\hat{\mathbf{L}}}^2 - \left\|\mathbf{e}^{(k+1)}\right\|_{\hat{\mathbf{L}}}^2 \ge 2\rho \left\|\mathbf{\Lambda}^2\left(\mathbf{\Sigma}^2 + \mu_k\mathbf{\Lambda}^2\right)^{-1}\mathbf{U}^T\mathbf{r}^{(k)}\right\|_2 \left\|\mathbf{r}^{(k)}\right\|_2.$$

Proof First, observe that

$$\mathbf{e}^{(k+1)} = \mathbf{y}^\dagger - \mathbf{y}^{(k+1)} = \mathbf{y}^\dagger - \mathbf{y}^{(k)} - \mathbf{h}^{(k)} = \mathbf{e}^{(k)} - \mathbf{h}^{(k)}.$$

Substituting the GSVD (16) into the definition of $\mathbf{h}^{(k)}$ in (13) gives

$$\mathbf{h}^{(k)} = \mathbf{X}^{-T}\left(\mathbf{\Sigma}^2 + \mu_k\mathbf{\Lambda}^2\right)^{-1}\mathbf{\Sigma}\mathbf{U}^T\mathbf{r}^{(k)}.$$

We now expand $\left\|\mathbf{e}^{(k+1)}\right\|_{\hat{\mathbf{L}}}^2$ as

$$\left\|\mathbf{e}^{(k+1)}\right\|_{\hat{\mathbf{L}}}^2 = \left\|\mathbf{e}^{(k)} - \mathbf{h}^{(k)}\right\|_{\hat{\mathbf{L}}}^2 = \left\|\mathbf{e}^{(k)}\right\|_{\hat{\mathbf{L}}}^2 - 2\left\langle\hat{\mathbf{L}}\mathbf{e}^{(k)}, \hat{\mathbf{L}}\mathbf{h}^{(k)}\right\rangle + \left\|\mathbf{h}^{(k)}\right\|_{\hat{\mathbf{L}}}^2.$$

Therefore, letting $\hat{\mathbf{r}}^{(k)} = \mathbf{U}^T \mathbf{r}^{(k)}$

$$
\begin{aligned}
&\left\|\mathbf{e}^{(k)}\right\|_{\hat{\mathbf{L}}}^2 - \left\|\mathbf{e}^{(k+1)}\right\|_{\hat{\mathbf{L}}}^2 = 2\left\langle \mathbf{\Lambda}\mathbf{X}^T\mathbf{e}^{(k)}, \mathbf{\Lambda}\mathbf{X}^T\mathbf{h}^{(k)}\right\rangle - \left\|\mathbf{h}^{(k)}\right\|_{\hat{\mathbf{L}}}^2 \\
&\geq 2\left\langle \mathbf{X}^T\mathbf{e}^{(k)}, \mathbf{\Lambda}^2\left(\mathbf{\Sigma}^2 + \mu_k\mathbf{\Lambda}^2\right)^{-1}\mathbf{\Sigma}\hat{\mathbf{r}}^{(k)}\right\rangle \\
&\quad - 2\left\langle \hat{\mathbf{r}}^{(k)}, \mathbf{\Lambda}^2\mathbf{\Sigma}^2\left(\mathbf{\Sigma}^2 + \mu_k\mathbf{\Lambda}^2\right)^{-2}\hat{\mathbf{r}}^{(k)}\right\rangle \\
&= 2\left\langle \hat{\mathbf{r}}^{(k)}, \mathbf{\Lambda}^2\left(\mathbf{\Sigma}^2 + \mu_k\mathbf{\Lambda}^2\right)^{-1}\hat{\mathbf{r}}^{(k)}\right\rangle \\
&\quad - 2\left\langle \hat{\mathbf{r}}^{(k)}, \mathbf{\Lambda}^2\mathbf{\Sigma}^2\left(\mathbf{\Sigma}^2 + \mu_k\mathbf{\Lambda}^2\right)^{-2}\hat{\mathbf{r}}^{(k)}\right\rangle \\
&\quad - 2\left\langle \hat{\mathbf{r}}^{(k)} - \mathbf{\Sigma}\mathbf{X}^T\mathbf{e}^{(k)}, \mathbf{\Lambda}^2\left(\mathbf{\Sigma}^2 + \mu_k\mathbf{\Lambda}^2\right)^{-1}\hat{\mathbf{r}}^{(k)}\right\rangle \\
&= 2\left\langle \hat{\mathbf{r}}^{(k)}, \mathbf{\Lambda}^2\left[\mu_k\mathbf{\Lambda}^2\left(\mathbf{\Sigma}^2 + \mu_k\mathbf{\Lambda}^2\right)^{-2}\right]\hat{\mathbf{r}}^{(k)}\right\rangle \\
&\quad - 2\left\langle \hat{\mathbf{r}}^{(k)} - \mathbf{\Sigma}\mathbf{X}^T\mathbf{e}^{(k)}, \mathbf{\Lambda}^2\left(\mathbf{\Sigma}^2 + \mu_k\mathbf{\Lambda}^2\right)^{-1}\hat{\mathbf{r}}^{(k)}\right\rangle \\
&\geq 2\mu_k\left\|\mathbf{\Lambda}^2\left(\mathbf{\Sigma}^2 + \mu_k\mathbf{\Lambda}^2\right)^{-1}\hat{\mathbf{r}}^{(k)}\right\|_2^2 \\
&\quad - 2\left\|\hat{\mathbf{r}}^{(k)} - \mathbf{\Sigma}\mathbf{X}^T\mathbf{e}^{(k)}\right\|_2\left\|\mathbf{\Lambda}^2\left(\mathbf{\Sigma}^2 + \mu_k\mathbf{\Lambda}^2\right)^{-1}\hat{\mathbf{r}}^{(k)}\right\|_2 \\
&= 2\left\|\mathbf{\Lambda}^2\left(\mathbf{\Sigma}^2 + \mu_k\mathbf{\Lambda}^2\right)^{-1}\hat{\mathbf{r}}^{(k)}\right\|_2\left(\left\|\mu_k\mathbf{\Lambda}^2\left(\mathbf{\Sigma}^2 + \mu_k\mathbf{\Lambda}^2\right)^{-1}\hat{\mathbf{r}}^{(k)}\right\|_2\right. \\
&\quad \left. - \left\|\hat{\mathbf{r}}^{(k)} - \mathbf{\Sigma}\mathbf{X}^T\mathbf{e}^{(k)}\right\|_2\right) \\
&\overset{(a)}{=} 2\left\|\mathbf{\Lambda}^2\left(\mathbf{\Sigma}^2 + \mu_k\mathbf{\Lambda}^2\right)^{-1}\hat{\mathbf{r}}^{(k)}\right\|_2\left(\left\|\mathbf{r}^{(k)} - \mathbf{B}_{p+1,p}\mathbf{h}^{(k)}\right\|_2 - \left\|\mathbf{r}^{(k)} - \mathbf{B}_{p+1,p}\mathbf{e}^{(k)}\right\|_2\right) \\
&\overset{(b)}{\geq} 2\left\|\mathbf{\Lambda}^2\left(\mathbf{\Sigma}^2 + \mu_k\mathbf{\Lambda}^2\right)^{-1}\hat{\mathbf{r}}^{(k)}\right\|_2\left(q_k\left\|\mathbf{r}^{(k)}\right\|_2 - \left(\rho + \frac{1+\rho}{\tau_k}\right)\left\|\mathbf{r}^{(k)}\right\|_2\right) \\
&\overset{(c)}{\geq} 2\left\|\mathbf{\Lambda}^2\left(\mathbf{\Sigma}^2 + \mu_k\mathbf{\Lambda}^2\right)^{-1}\hat{\mathbf{r}}^{(k)}\right\|_2\left(\left(2\rho + \frac{1+\rho}{\tau_k}\right)\left\|\mathbf{r}^{(k)}\right\|_2\right. \\
&\quad \left. - \left(\rho + \frac{1+\rho}{\tau_k}\right)\left\|\mathbf{r}^{(k)}\right\|_2\right) \\
&= 2\rho\left\|\mathbf{\Lambda}^2\left(\mathbf{\Sigma}^2 + \mu_k\mathbf{\Lambda}^2\right)^{-1}\hat{\mathbf{r}}^{(k)}\right\|_2\left\|\mathbf{r}^{(k)}\right\|_2,
\end{aligned}
$$

where (a) follows from (18), (b) is a consequence of the definition of μ_k in (14) and Lemma 1, and (c) holds due to the definition of q_k in (12). □

Proposition 2 yields the following result.

Corollary 3 *With the notation and assumptions above, we have*

$$\left\|\mathbf{e}^{(0)}\right\|_{\hat{\mathbf{L}}}^2 \geq 2\rho \sum_{k=0}^{k^\delta-1} \left\|\boldsymbol{\Lambda}^2\left(\boldsymbol{\Sigma}^2+\mu_k\boldsymbol{\Lambda}^2\right)^{-1}\mathbf{U}^T\mathbf{r}^{(k)}\right\|_2 \left\|\mathbf{r}^{(k)}\right\|_2 \geq c\sum_{k=0}^{k^\delta-1}\left\|\mathbf{r}^{(k)}\right\|_2^2,$$

where c is constant that depends only on ρ and q.

Proof The proof is similar to the proof of [13, Corollary 3]. We therefore omit the details. □

The immediate implication of Corollary 3 is that the proposed iterative solution method terminates after a finite number of iterations if $\delta > 0$. In fact, if this were not the case, then we would have that $k^\delta = \infty$ and the series

$$\sum_{k=0}^{\infty}\left\|\mathbf{r}^{(k)}\right\|_2$$

would converge. Therefore,

$$\lim_{k\to\infty}\left\|\mathbf{r}^{(k)}\right\|_2 = 0,$$

which is absurd since we assumed that $\left\|\mathbf{r}^{(k)}\right\|_2 \geq \tau\delta > 0$.

We now show that, if $\delta = 0$, then the iterates (15) converge to a solution of $\mathbf{A}^T\mathbf{y} = \widetilde{\mathbf{b}}$ as k increases.

Theorem 4 *Let $\delta = 0$ and assume that $\mathbf{B}_{p+1,p}\mathbf{y}^{(0)} \neq \widetilde{\mathbf{b}}$. Then the sequence $\{\mathbf{y}^{(k)}\}_k$ converges to a solution of the system $\mathbf{B}_{p+1,p}\mathbf{y} = \widetilde{\mathbf{b}}$ that depends on $\mathbf{y}^{(0)}$.*

Proof We first show that if the iterates $\{\mathbf{y}^{(k)}\}_k$ converge, then this requires an infinite number of iterations. By contradiction, assume that for a finite k it holds $\mathbf{B}_{p+1,p}\mathbf{y}^{(k)} = \widetilde{\mathbf{b}}$. Then $\mathbf{h}^{(k-1)}$ coincides with $\mathbf{e}^{(k-1)}$ (up to a component in the null space of $\mathbf{B}_{p+1,p}$). This implies

$$\begin{aligned} q_{k-1}\left\|\mathbf{r}^{(k-1)}\right\|_2 &= \left\|\mathbf{r}^{(k-1)} - \mathbf{B}_{p+1,p}\mathbf{h}^{(k-1)}\right\|_2 = \left\|\mathbf{r}^{(k-1)} - \mathbf{B}_{p+1,p}\mathbf{e}^{(k-1)}\right\|_2 \\ &\leq \left(\rho + \frac{1-\rho}{\tau_{k-1}}\right)\left\|\mathbf{r}^{(k-1)}\right\|_2. \end{aligned}$$

This, however, contradicts the definition of q_{k-1} in (12). Therefore, we must have that there is no finite k such that $\mathbf{B}_{p+1,p}\mathbf{y}^{(k)} = \widetilde{\mathbf{b}}$.

We now show that $\{\mathbf{y}^{(k)}\}_k$ is a Cauchy sequence. Let $k > l$. Then

$$\begin{aligned} \left\|\mathbf{y}^{(k)} - \mathbf{y}^{(l)}\right\|_{\hat{\mathbf{L}}}^2 = \left\|\mathbf{e}^{(k)} - \mathbf{e}^{(l)}\right\|_{\hat{\mathbf{L}}}^2 &= \left\|\mathbf{e}^{(k)}\right\|_{\hat{\mathbf{L}}}^2 - 2\left\langle \hat{\mathbf{L}}\mathbf{e}^{(k)}, \hat{\mathbf{L}}\mathbf{e}^{(l)}\right\rangle + \left\|\mathbf{e}^{(l)}\right\|_{\hat{\mathbf{L}}}^2 \\ &= \left\|\mathbf{e}^{(k)}\right\|_{\hat{\mathbf{L}}}^2 - \left\|\mathbf{e}^{(l)}\right\|_{\hat{\mathbf{L}}}^2 - 2\left\langle \hat{\mathbf{L}}\mathbf{e}^{(k)} - \hat{\mathbf{L}}\mathbf{e}^{(l)}, \hat{\mathbf{L}}\mathbf{e}^{(l)}\right\rangle \\ &= \left\|\mathbf{e}^{(k)}\right\|_{\hat{\mathbf{L}}}^2 - \left\|\mathbf{e}^{(l)}\right\|_{\hat{\mathbf{L}}}^2 + 2\left\langle \hat{\mathbf{L}}\mathbf{e}^{(l)}, \hat{\mathbf{L}}\mathbf{y}^{(k)} - \hat{\mathbf{L}}\mathbf{y}^{(l)}\right\rangle. \end{aligned}$$

Therefore, it holds

$$\begin{aligned}
\left\|\mathbf{y}^{(k)}-\mathbf{y}^{(l)}\right\|_{\hat{\mathbf{L}}}^{2} &= \left\|\mathbf{e}^{(k)}\right\|_{\hat{\mathbf{L}}}^{2}-\left\|\mathbf{e}^{(l)}\right\|_{\hat{\mathbf{L}}}^{2}+2\sum_{i=l}^{k-1}\left\langle \hat{\mathbf{L}}\mathbf{e}^{(l)}, \boldsymbol{\Lambda}\mathbf{X}^{T}\mathbf{h}^{(i)}\right\rangle \\
&= \left\|\mathbf{e}^{(k)}\right\|_{\hat{\mathbf{L}}}^{2}-\left\|\mathbf{e}^{(l)}\right\|_{\hat{\mathbf{L}}}^{2}+2\sum_{i=l}^{k-1}\left\langle \boldsymbol{\Lambda}\mathbf{X}^{T}\mathbf{e}^{(l)}, \boldsymbol{\Lambda}\left(\boldsymbol{\Sigma}^{2}+\mu_{i}\boldsymbol{\Lambda}^{2}\right)^{-1}\boldsymbol{\Sigma}\mathbf{U}^{T}\mathbf{r}^{(i)}\right\rangle \\
&= \left\|\mathbf{e}^{(k)}\right\|_{\hat{\mathbf{L}}}^{2}-\left\|\mathbf{e}^{(l)}\right\|_{\hat{\mathbf{L}}}^{2}+2\sum_{i=l}^{k-1}\left\langle \boldsymbol{\Sigma}X\mathbf{b}^{T}\mathbf{e}^{(l)}, \boldsymbol{\Lambda}^{2}\left(\boldsymbol{\Sigma}^{2}+\mu_{i}\boldsymbol{\Lambda}^{2}\right)^{-1}\mathbf{U}^{T}\mathbf{r}^{(i)}\right\rangle \\
&\leq \left\|\mathbf{e}^{(k)}\right\|_{\hat{\mathbf{L}}}^{2}-\left\|\mathbf{e}^{(l)}\right\|_{\hat{\mathbf{L}}}^{2}+2\sum_{i=l}^{k-1}\left\|\boldsymbol{\Sigma}\mathbf{X}^{T}\mathbf{e}^{(l)}\right\|_{2}\left\|\boldsymbol{\Lambda}^{2}\left(\boldsymbol{\Sigma}^{2}+\mu_{i}\boldsymbol{\Lambda}^{2}\right)^{-1}\mathbf{U}^{T}\mathbf{r}^{(i)}\right\|_{2} \\
&= \left\|\mathbf{e}^{(k)}\right\|_{\hat{\mathbf{L}}}^{2}-\left\|\mathbf{e}^{(l)}\right\|_{\hat{\mathbf{L}}}^{2}+2\sum_{i=l}^{k-1}\left\|\mathbf{B}_{p+1,p}\mathbf{e}^{(l)}\right\|_{2}\left\|\boldsymbol{\Lambda}^{2}\left(\boldsymbol{\Sigma}^{2}+\mu_{i}\boldsymbol{\Lambda}^{2}\right)^{-1}\mathbf{U}^{T}\mathbf{r}^{(i)}\right\|_{2} \\
&= \left\|\mathbf{e}^{(k)}\right\|_{\hat{\mathbf{L}}}^{2}-\left\|\mathbf{e}^{(l)}\right\|_{\hat{\mathbf{L}}}^{2}+2\sum_{i=l}^{k-1}\left\|\mathbf{r}^{(l)}\right\|_{2}\left\|\boldsymbol{\Lambda}^{2}\left(\boldsymbol{\Sigma}^{2}+\mu_{i}\boldsymbol{\Lambda}^{2}\right)^{-1}\mathbf{U}^{T}\mathbf{r}^{(i)}\right\|_{2}.
\end{aligned}$$

Similarly, let $j < l$. Then

$$\begin{aligned}
\left\|\mathbf{y}^{(l)}-\mathbf{y}^{(j)}\right\|_{\hat{\mathbf{L}}}^{2} &= \left\|\mathbf{e}^{(l)}\right\|_{\hat{\mathbf{L}}}^{2}-\left\|\mathbf{e}^{(j)}\right\|_{\hat{\mathbf{L}}}^{2}+2\sum_{i=j}^{j-1}\left\langle \hat{\mathbf{L}}\mathbf{e}^{(l)}, \boldsymbol{\Lambda}\mathbf{X}^{T}\mathbf{h}^{(i)}\right\rangle \\
&= \left\|\mathbf{e}^{(l)}\right\|_{\hat{\mathbf{L}}}^{2}-\left\|\mathbf{e}^{(j)}\right\|_{\hat{\mathbf{L}}}^{2}+2\sum_{i=j}^{l-1}\left\langle \boldsymbol{\Lambda}\mathbf{X}^{T}\mathbf{e}^{(l)}, \boldsymbol{\Lambda}\left(\boldsymbol{\Sigma}^{2}+\mu_{i}\boldsymbol{\Lambda}^{2}\right)^{-1}\boldsymbol{\Sigma}\mathbf{U}^{T}\mathbf{r}^{(i)}\right\rangle \\
&= \left\|\mathbf{e}^{(l)}\right\|_{\hat{\mathbf{L}}}^{2}-\left\|\mathbf{e}^{(j)}\right\|_{\hat{\mathbf{L}}}^{2}+2\sum_{i=j}^{l-1}\left\langle \boldsymbol{\Sigma}\mathbf{X}^{T}\mathbf{e}^{(l)}, \boldsymbol{\Lambda}^{2}\left(\boldsymbol{\Sigma}^{2}+\mu_{i}\boldsymbol{\Lambda}^{2}\right)^{-1}\mathbf{U}^{T}\mathbf{r}^{(i)}\right\rangle \\
&\leq \left\|\mathbf{e}^{(l)}\right\|_{\hat{\mathbf{L}}}^{2}-\left\|\mathbf{e}^{(j)}\right\|_{\hat{\mathbf{L}}}^{2}+2\sum_{i=j}^{l-1}\left\|\boldsymbol{\Sigma}\mathbf{X}^{T}\mathbf{e}^{(l)}\right\|_{2}\left\|\boldsymbol{\Lambda}^{2}\left(\boldsymbol{\Sigma}^{2}+\mu_{i}\boldsymbol{\Lambda}^{2}\right)^{-1}\mathbf{U}^{T}\mathbf{r}^{(i)}\right\|_{2} \\
&= \left\|\mathbf{e}^{(l)}\right\|_{\hat{\mathbf{L}}}^{2}-\left\|\mathbf{e}^{(j)}\right\|_{\hat{\mathbf{L}}}^{2}+2\sum_{i=j}^{l-1}\left\|\mathbf{B}_{p+1,p}\mathbf{e}^{(l)}\right\|_{2}\left\|\boldsymbol{\Lambda}^{2}\left(\boldsymbol{\Sigma}^{2}+\mu_{i}\boldsymbol{\Lambda}^{2}\right)^{-1}\mathbf{U}^{T}\mathbf{r}^{(i)}\right\|_{2} \\
&= \left\|\mathbf{e}^{(l)}\right\|_{\hat{\mathbf{L}}}^{2}-\left\|\mathbf{e}^{(j)}\right\|_{\hat{\mathbf{L}}}^{2}+2\sum_{i=j}^{l-1}\left\|\mathbf{r}^{(l)}\right\|_{2}\left\|\boldsymbol{\Lambda}^{2}\left(\boldsymbol{\Sigma}^{2}+\mu_{i}\boldsymbol{\Lambda}^{2}\right)^{-1}\mathbf{U}^{T}\mathbf{r}^{(i)}\right\|_{2}.
\end{aligned}$$

We obtain

$$\begin{aligned}\left\|\mathbf{y}^{(k)}-\mathbf{y}^{(j)}\right\|_{\hat{\mathbf{L}}}^{2} &\leq 2\left\|\mathbf{y}^{(k)}-\mathbf{y}^{(l)}\right\|_{\hat{\mathbf{L}}}^{2}+2\left\|\mathbf{y}^{(l)}-\mathbf{y}^{(j)}\right\|_{\hat{\mathbf{L}}}^{2} \\ &\leq 2\left\|\mathbf{e}^{(k)}\right\|_{\hat{\mathbf{L}}}^{2}+2\left\|\mathbf{e}^{(j)}\right\|_{\hat{\mathbf{L}}}^{2}-4\left\|\mathbf{e}^{(l)}\right\|_{\hat{\mathbf{L}}}^{2} \\ &\quad +4\sum_{i=j}^{k-1}\left\|\mathbf{r}^{(l)}\right\|_{2}\left\|\mathbf{\Lambda}^{2}\left(\mathbf{\Sigma}^{2}+\mu_i\mathbf{\Lambda}^{2}\right)^{-1}\mathbf{U}^{T}\mathbf{r}^{(i)}\right\|_{2}.\end{aligned}$$

Choosing $j \in \{l, \ldots, k-1\}$ such that $\left\|\mathbf{r}^{(j)}\right\|_2$ is minimal yields

$$\begin{aligned}\left\|\mathbf{y}^{(k)}-\mathbf{y}^{(j)}\right\|_{\hat{\mathbf{L}}}^{2} &\leq 2\left\|\mathbf{e}^{(k)}\right\|_{\hat{\mathbf{L}}}^{2}+2\left\|\mathbf{e}^{(j)}\right\|_{\hat{\mathbf{L}}}^{2}-4\left\|\mathbf{e}^{(l)}\right\|_{\hat{\mathbf{L}}}^{2} \\ &\quad +4\sum_{i=j}^{k-1}\left\|\mathbf{r}^{(i)}\right\|_{2}\left\|\mathbf{\Lambda}^{2}\left(\mathbf{\Sigma}^{2}+\mu_i\mathbf{\Lambda}^{2}\right)^{-1}\mathbf{U}^{T}\mathbf{r}^{(i)}\right\|_{2}.\end{aligned}$$

Since the sequence $\left\{\left\|\mathbf{e}^{(k)}\right\|_{\hat{\mathbf{L}}}^{2}\right\}_k$ is monotonically decreasing and bounded from below, it is convergent and, if $k, j \to \infty$, the quantity

$$\sum_{i=j}^{k-1}\left\|\mathbf{r}^{(i)}\right\|_{2}\left\|\mathbf{\Lambda}^{2}\left(\mathbf{\Sigma}^{2}+\mu_i\mathbf{\Lambda}^{2}\right)^{-1}\mathbf{U}^{T}\mathbf{r}^{(i)}\right\|_{2}$$

is the tail of a convergent series thanks to Corollary 3. Therefore, the right-hand side of the inequality above goes to zero when $k, j \to \infty$. This shows that the sequence $\left\{\mathbf{y}^{(k)}\right\}_k$ converges.

It follows from Corollary 3 that the series

$$\sum_{k=1}^{\infty}\left\|\mathbf{r}^{(k)}\right\|_{2}^{2}$$

converges and, therefore,

$$\lim_{k\to\infty}\left\|\mathbf{r}^{(k)}\right\|_{2}^{2}=0.$$

This implies that the limit point of the sequence $\left\{\mathbf{y}^{(k)}\right\}_k$ is a solution of $\mathbf{B}_{p+1,p}\mathbf{y}=\mathbf{b}$.□

We now will show that when the norm of the noise in $\mathbf{b}^{\delta}$ goes to 0, the approximate solutions generated by our method converge to a solution of the noise-free problem.

Theorem 5 *Let $\delta \mapsto \widetilde{\mathbf{b}}^{\delta}$ be a function such that*

$$\left\|\widetilde{\mathbf{b}}^{\delta}-\widetilde{\mathbf{b}}\right\|_2 \leq \delta \quad \forall\, \delta > 0.$$

Fix ρ, τ, q, and $\mathbf{y}^{(0)}$. Let k^δ denote the number of iterations required to satisfy the stopping criterion for noise level δ, and let $\mathbf{y}^\delta$ be the corresponding computed approximate solution. Then, for $\delta \to 0$, we have that $\mathbf{y}^\delta$ converges to a solution of $\mathbf{B}_{p+1,p}\mathbf{y} = \mathbf{b}$ that depends on $\mathbf{y}^{(0)}$.

Proof We omit the proof since it is similar to the proof of [20, Theorem 2.3]. □

4 Computed examples

We compare several versions of the PIT-A and PIT-GKB methods described in the previous sections to various other solution methods for (3) and (5), that are summarized in Table 1.

Apart from methods that implement iterated Tikhonov regularization through projection onto Krylov subspaces, we consider the AT and GKBT methods that project a Tikhonov minimization problem onto Krylov subspaces of increasing dimension generated by the Arnoldi and Golub-Kahan processes, respectively. The AT and GKBT methods can be regarded as straightforward simplifications of the PIT-A and PIT-GKB methods, respectively. Since AT-sft and GKBT-sft are the AT and GKBT methods applied to a Tikhonov problem transformed into standard form, respectively, $\mathbf{L}_{\mathbf{A}}^{\dagger}$ acts as a preconditioner, or "priorconditioner" [21], for the solution subspace generated by the Arnoldi and Golub-Kahan processes, respectively. We also consider the TA and TGKB methods, which compute approximate solutions of (2) by truncated iteration applied to the Arnoldi and Golub-Kahan processes and correspond to early stopping of the GMRES and LSQR methods, respectively. In the TA method, the Arnoldi process is applied to reduce the matrix $\mathbf{A}$ with initial vector $\mathbf{b}^\delta$ (cf. (8)). The solution of the reduced problem is an approximation of the solution of (2). The Arnoldi process is terminated at step k with k chosen as small as possible so that the computed approximate solution $\mathbf{x}_k$ of (2) satisfies the discrepancy principle

$$\left\|\mathbf{A}\mathbf{x}_k - \mathbf{b}^\delta\right\|_2 \leq \tau\delta.$$

The TGKB method is defined analogously.

Table 1 (Acronyms of the) projection methods considered in the numerical tests

TA, TGKB	truncated Arnoldi, Golub-Kahan-Bidiagonalization (resp.) process
AT, GKBT	Arnoldi, Golub-Kahan-Bidiagonalization (resp.) Tikhonov
AT-sft, GKBT-sft	AT, GKBT (resp.) for standard form transformed Tikhonov
AT-L GKBT-L	AT, GKBT (resp.) with a general regularization matrix (11)
PIT	Projected Iterated Tikhonov
PIT-A, PIT-GKB	Arnoldi, Golub-Kahan-Bidiagonalization (resp.) based PIT
PIT-A-sf, PIT-GKB-sf	PIT-A, PIT-GKB (resp.) for Tikhonov in standard form
PIT-A-sft, PIT-GKB-sft	PIT-A, PIT-GKB (resp.) for standard form transformed Tikhonov
PNITR-GKB	Projected Iterated Tikhonov [9]

When applying the discrepancy principle, one typically chooses the regularization parameter $\mu > 0$ so that the solution $\mathbf{x}_\mu$ of the Tikhonov minimization problem (3) satisfies

$$\|\mathbf{A}\mathbf{x}_\mu - \mathbf{b}^\delta\|_2 = \tau\delta, \tag{19}$$

where $\delta \geq 0$ is the bound in (6) and $\tau > 1$, $\tau \simeq 1$, is a safety parameter. This is motivated by the fact that, if $\mathbf{A}\mathbf{x}^\dagger = \mathbf{b}$, then

$$\|\mathbf{A}\mathbf{x}^\dagger - \mathbf{b}^\delta\|_2 = \|\mathbf{b} - \mathbf{b}^\delta\|_2 \leq \delta.$$

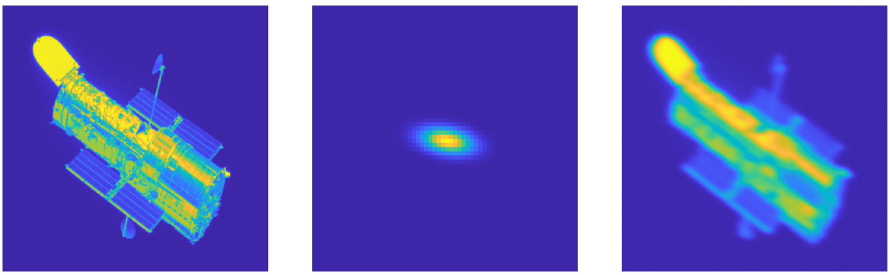
To assess the quality of the reconstructed solutions, we evaluate the Relative Reconstruction Error (RRE) for the reconstruction $\mathbf{x}^{(k)}$. It is defined as

$$\text{RRE} := \text{RRE}(\mathbf{x}^{(k)}, \mathbf{x}^\dagger) = \frac{\left\|\mathbf{x}^{(k)} - \mathbf{x}^\dagger\right\|_2}{\left\|\mathbf{x}^\dagger\right\|_2}. \tag{20}$$

For all test cases, we perturb the measurements with white Gaussian noise, i.e., the noise vector $\mathbf{e}$ is drawn from a Gaussian with zero mean and a scaled identity covariance matrix; we refer to the ratio $\sigma = \|\mathbf{e}\|_2/\|\mathbf{A}\mathbf{x}\|_2$ as the noise level.

In our numerical experiments, we choose $q = 0.7$ and $\rho = 10^{-3}$ for all PIT methods (as well as AIT variants). Additionally, we let $\tau = 1.01$ when determining the regularization parameters μ_k and using the discrepancy principle as stopping criterion.

Example 1 - image deblurring We consider an image deblurring test problem. The true image is shown in Fig. 1a. It is made up of 512×512 pixels and is available from IR Tools [22]. The true image is corrupted with an anisotropic Gaussian point spread function (PSF) of dimension 64×64, which is shown in Fig. 1b. The blurring operator $\mathbf{A} \in \mathbb{R}^{262144 \times 262144}$ is defined implicitly by the PSF. The observed blurred and noise-contaminated image with noise level $\sigma = 1\%$ is shown in Fig. 1c. The noise- and blur-free image, which is assumed not to be available, is represented by the vector $\mathbf{x}^\dagger \in \mathbb{R}^{512^2}$, while the available blur- and noise-contaminated image is represented by the vector $\mathbf{b}^\delta \in \mathbb{R}^{512^2}$.

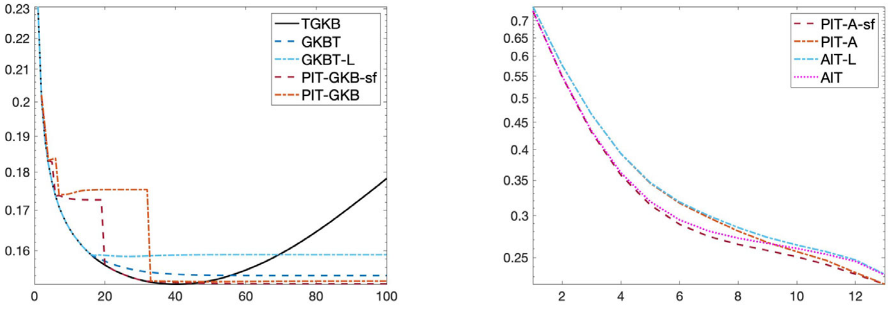


(a) (b) (c)

Fig. 1 Image deblurring test problem. (a) True image of size 512×512. (b) Anisotropic Gaussian PSF, and (c) Blurred and noisy image with 1% Gaussian noise

The aim of this test problem is to compare the quality of the reconstructions obtained with the Arnoldi and Golub-Kahan-based methods proposed in Section 2. In particular, we consider PIT-A-sf and PIT-A, which both use the Arnoldi process, and PIT-GKB and PIT-GKB-sf, which both use Golub-Kahan bidiagonaliztion; Tikhonov regularization is considered both in standard and general form. We compare our methods to the *approximated iterated Tikhonov* method developed by Donatelli-Hanke [13] that we denote by AIT (and AIT-L when $L \neq I$). In addition, we consider the TGKB, GKBT, and GKBT-L methods. For all solvers, we apply a maximum of 100 steps of the Arnoldi and Golub-Kahan bidiagonlization processes. The convergence history for 100 iterations for all methods considered is displayed in Fig. 2. On the left panel, we present the RRE of methods based on Golub-Kahan. On the right panel we show RRE convergence history versus the number of Iterated Tikhonov iterations for the projected problem of order $p = 13$ (i.e., at Arnoldi iteration number 13) for PIT-A, PIT-A-sf, and for AIT and AIT-L when they satisfy the discrepancy stopping criteria (19) (13 iterations). Moreover, to numerically study the robustness of the methods to noise, we experiment with two noise levels: 1% and 5%. Results are summarized in Table 2. Finally, we presented the reconstructed images in Fig. 3. We observe that the methods that we propose (both when they are based on the Arnoldi and Golub-Kahan processes) provide restorations of higher quality than the other methods in our comparison. The TGKB method semiconverges, while all other methods that apply Tikhonov regularization and rely on either the Arnold or Golub-Kahan processes yield a stabilized solution and improve the reconstruction quality with each iteration.

Example 2 - tomography problem This test problem is generated by using the AIR Tools II [23] functionalities available within IR Tools. Specifically, we consider a smooth phantom of size 256×256 pixels (shown in Fig. 6 (a)) and use an under-sampled fan beam acquisition setting, whereby only one of every four angles between 0° and 180° is sampled. We add 10% white Gaussian noise to the exact data. The associated discrete forward operator is represented by a matrix $\mathbf{A}$ of size 17014×65536.



(a) (b)

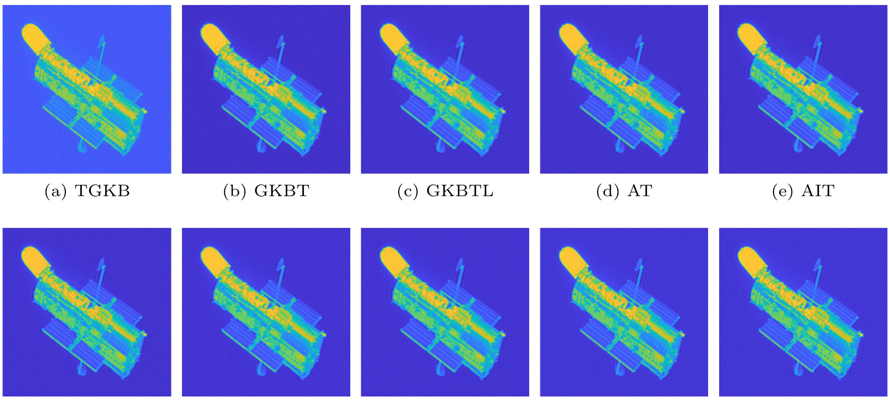
Fig. 2 Image deblurring test problem. Left: RRE convergence history for methods based on Golub-Kahan such as TGKB, GKBT, GKBT-L, PIT-GKB-sf, and PIT-GKB. Right: RRE convergence history versus the number of Iterated Tikhonov iterations for the projected problem of order p = 13 (i.e., at Arnoldi iteration number) for PIT-A, PIT-A-sf, and for AIT and AIT-L when they satisfy the discrepancy stopping criteria (19) (13 iterations)

Table 2 Image deblurring test problem. RRE for noise levels 1% and 5% for all methods

σ	TGKB	GKBT	GKBT-L	AT	AIT
1%	0.1787	0.1543	0.1591	0.1543	0.1745
5%	0.1776	0.1833	0.2149	0.1833	0.1745
σ	AIT-L	PIT-A	PIT-A-sf	PIT-GKB	PIT-GKB-sf
1%	0.1750	0.1573	0.1566	0.1521	0.1521
5%	0.1750	0.1574	0.1567	0.1523	0.1523

Since the matrix A is rectangular, we cannot use reduction techniques based on the Arnoldi process. We therefore compare the reconstructions obtained by methods based on partial Golub-Kahan bidiagonalization (GKB). For the GKBT-sft, GKBT-L, PNITR-GKB, PIT-GKB, and PIT-GKB-sft methods, the regularization matrix is chosen to be a rescaled 2D discretized Laplacian with periodic boundary conditions; see, e.g., [24]. For all methods we apply a maximum of 30 steps of the GKB process to the matrix $\mathbf{A}$ with initial vector $\mathbf{b}^\delta$. Figure 4 displays the history of the relative errors versus the dimension of the GKB-generated solution subspaces for each solution method.

Figure 4 shows the new iterated Tikhonov methods PIT-GKB-sf and PIT-GKB determine approximate solutions of higher quality than their non-iterated Tikhonov counterparts, with PIT-GKB delivering the best reconstructions. Indeed, after the discrepancy principle is satisfied at the 3rd GKB iteration (so that $k_{DP} = 3$ in (7)), TGKB abruptly semiconverges, while all the methods that apply some form of Tikhonov regularization start to automatically choose an iteration-dependent Tikhonov regularization parameter, thereby stabilizing and possibly improving the quality of the computed solution. In frame 4a we see that GKBT-sft, based on 'priorconditioned' GKB, converges slower than all the other methods based on GKB; this behavior motivates the application of (iterated) Tikhonov regularization in general form to the projected prob-



(a) TGKB (b) GKBT (c) GKBTL (d) AT (e) AIT

(f) AIT-L (g) PIT-A (h) PIT-A-sf (i) PIT-GKB (j) PIT-GKB-sf

Fig. 3 Image deblurring test problem. First row: reconstructions with TGKB, GKBT, GKBT-L, AT, AIT; Second row: reconstructions with AIT-L, PIT-A, PIT-A-sf, PIT-GKB, and PIT-GKB-sf

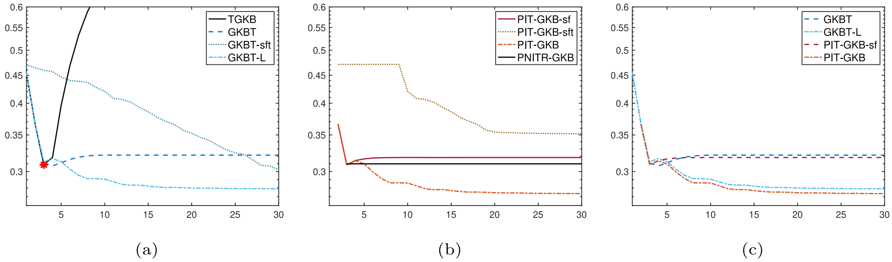


Fig. 4 Tomography test problem. History of the relative errors versus the dimension of the GKB-generated approximation subspace for different solvers. (a) purely iterative methods and hybrid methods. (b) methods that apply iterated Tikhonov on the projected problem. (c) closer comparison of some of the hybrid methods and iterated Tikhonov methods. A red asterisk in frame (a) highlights the GKB iteration satisfying the discrepancy principle

lem generated by GKB, rather than adopting available (iterated) Tikhonov in standard form to the projected problem generated by 'priorconditioned' GKB, i.e., associated to a Tikhonov problem in general form transformed into standard form (see also frame 4b). Also, looking at Fig. 4c, it is evident that incorporating a regularization matrix positively impacts the quality of the reconstruction, as measured by the relative reconstruction error, although this is true already for GKBT-L (when compared to GKBT), further improvements are possible when applying PIT-GKB.

In general, looking at frames 4a and c we see that, applying some additional (either Tikhonov or iterated Tikhonov) regularization to the projected problem stabilizes the behavior of the reconstruction algorithm (with the general form of the latter delivering the most improved and stable results). Looking more closely at the behavior of the projected iterated Tikhonov methods in frame (b), we also see that the relative errors of the reconstructions delivered by PNITR-GKB rapidly stagnate far away from the PIT-GKB-sf and PIT-GKB ones. The behavior of the iterated Tikhonov iterations for PIT-GKB-sf and PIT-GKB applied to projected problems of different dimensions are displayed in Fig. 5. Here, we clearly see that, although iterated Tikhonov in standard form seems to perform slightly better than its general form counterpart during the first iterated Tikhonov iterations, at convergence the latter is always gives at least as

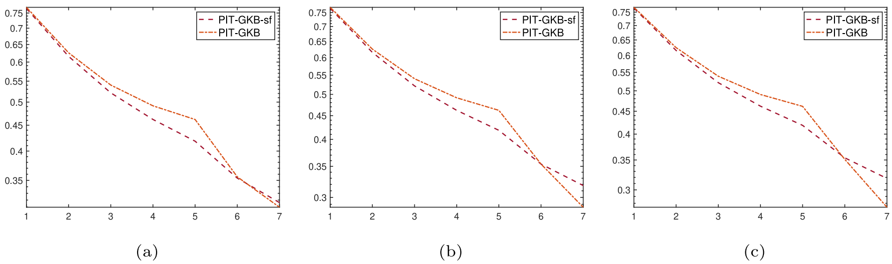


Fig. 5 Tomography test problem. History of the relative errors versus number of Iterated Tikhonov iterations for projected problems of order (i.e., at the GKB iteration number): (a) 5; (b) 10; (c) 20

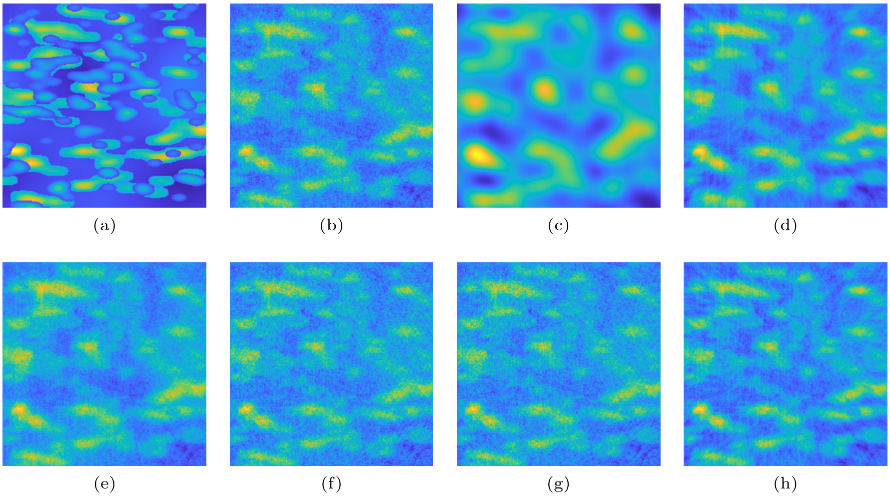


Fig. 6 Tomography test problem. (a) exact `threephasessmooth` phantom from AIR Tools II [23]. Reconstructions at: (b) the 4th TGKB iteration; (c) the 20th TGKB-sft iteration; (d) the 20th GKBT-L iteration; (e) the 20th PNITR-GKB iteration; (f) the 20th PIT-GKB-sf iteration (7th iterated Tikhonov iteration); (g) the 20th PIT-GKB-sft iteration (7th iterated Tikhonov iteration); (h) the 20th PIT-GKB iteration (7th iterated Tikhonov iteration)

accurate restorations as the former. This is especially true for projection subspaces of larger dimensions.

Finally, Fig. 6 shows some reconstructed phantoms. Here, we can see that the reconstructions associated to 'priorconditioned' GKB are over-smoothed; that is, more GKB iterations may be necessary to generate regularized solutions of high quality. There is not much visual difference between the reconstructions obtained by methods that combine Tikhonov and GKB projection. Arguably, some portions (especially the yellow spots) of the reconstruction obtained by PIT-GKB are more defined than for all the other reconstructions, and there is less smearing of the phantom structures.

5 Conclusion

This paper discusses iterated projected Tikhonov regularization. The regularized problems are projected into fairly low-dimensional Krylov subspaces that are determined by carrying out a small number of steps of the Arnoldi process or Golub-Kahan bidiagonalization. The regularization parameter is determined by using an extension of a technique introduced by Donatelli and Hanke [13]. Fairly general regularization matrices are allowed and a convergence analysis is provided. Computed examples show that the proposed methods are competitive with related Krylov subspace methods in terms of accuracy.

Acknowledgements AB and SG are members of the GNCS group of INdAM.

Author Contributions All the authors contributed equally to this work.

Funding Open access funding provided by Università degli Studi di Cagliari within the CRUI-CARE Agreement. AB is partially supported by INdAM-GNCS 2024 Project "Algebra lineare numerica per problemi di grandi dimensioni: aspetti teorici e applicazioni" (CUP_E53C23001670001). AB 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). AB 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). LO is partially supported by the U.S. National Science Foundation, Grants DMS-2038118 and DMS-2208294. Both SG and MP acknowledge the Isaac Newton Institute for Mathematical Sciences, Cambridge, for the support and hospitality during the programme "Rich and Nonlinear Tomography - a multidisciplinary approach" (supported by EPSRC grant no EP/R014604/) where partial work on this paper was undertaken. MP acknowledges partial support from NSF MSPRF Award No. 2202846 and NSF DMS No. 2410699.

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

Materials Availability Not applicable.

Code Availability The code is available upon request to the authors.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

Competing interests The authors declare no competing interests.

Ethics approval and consent to participate Not applicable.

Consent for publication Not applicable.

References

1. Chung, J., Gazzola, S.: Computational methods for large-scale inverse problems: a survey on hybrid projection methods. SIAM Rev. **66**, 205–284 (2024)
2. Bianchi, D., Buccini, A., Donatelli, M., Serra-Capizzano, S.: Iterated fractional Tikhonov regularization. Inverse Prob. **31**(5), 055005 (2015)
3. Buccini, A.: Regularizing preconditioners by non-stationary iterated Tikhonov with general penalty term. Appl. Numer. Math. **116**, 64–81 (2017)
4. Buccini, A., Donatelli, M., Reichel, L.: Iterated Tikhonov with general penalty term. Numer. Linear Algebra Appl. **24**, 2089 (2017)

5. Buccini, A., Donatelli, M., Reichel, L.: Theoretical and numerical aspects of a non-stationary preconditioned iterative method for linear discrete ill-posed problems. J. Comput. Appl. Math. **423**, 114940 (2023)
6. Buccini, A., Onisk, L., Reichel, L.: An Arnoldi-based preconditioner for iterated Tikhonov regularization. Numer. Algo. **92**(1), 223–245 (2023)
7. Engl, H.W., Hanke, M., Neubauer, A.: Regularization of Inverse Problems. Kluwer, Dordrecht (1996)
8. Huang, G., Reichel, L., Yin, F.: Projected nonstationary iterated Tikhonov regularization. BIT Numer. Math. **56**(2), 467–487 (2016)
9. Huang, G., Reichel, L., Yin, F.: On the choice of solution subspace for nonstationary iterated Tikhonov regularization. Numer. Algo. **72**, 1043–1063 (2016)
10. 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)
11. Kindermann, S., Raik, K.: A simplified L-curve as error estimator. Electron. Trans. Numer. Anal. **53**, 213–238 (2020)
12. Reichel, L., Rodriguez, G.: Old and new parameter choice rules for discrete ill-posed problems. Numer. Algo. **63**, 65–87 (2013)
13. Donatelli, M., Hanke, M.: Fast nonstationary preconditioned iterative methods for ill-posed problems, with application to image deblurring. Inverse Prob. **29**(9), 095008 (2013)
14. Buccini, A., Donatelli, M., Reichel, L., Zhang, W.: A new non-stationary preconditioned iterative method for linear discrete ill-posed problems. Numer. Linear Algebra Appl. **28**(2), 2353 (2021)
15. Golub, G.H., Van Loan, C.F.: Matrix Computations, 4th edn. Johns Hopkins University Press, Baltimore (2013)
16. Saad, Y.: Iterative Methods for Sparse Linear Systems, 2nd edn. SIAM, Philadelphia (2003)
17. Hochstenbach, M.E., Reichel, L.: An iterative method for Tikhonov regularization with a general linear regularization operator. J. Integral Equ. Appl. **22**, 463–480 (2010)
18. Huang, G., Reichel, L., Yin, F.: On the choice of solution subspace for large-scale Tikhonov regularization problems in general form. Numer. Algo. **81**, 33–55 (2019)
19. Dykes, L., Noschese, S., Reichel, L.: Rescaling the GSVD with application to ill-posed problems. Numer. Algo. **68**, 531–545 (2015)
20. Hanke, M.: A regularizing Levenberg-Marquardt scheme, with applications to inverse groundwater filtration problems. Inverse Prob. **13**(1), 79 (1997)
21. Calvetti, D., Somersalo, E.: Priorconditioners for linear systems. Inverse Prob. **21**(4), 1397 (2005)
22. Gazzola, S., Hansen, P.C., Nagy, J.G.: IR Tools: a MATLAB package of iterative regularization methods and large-scale test problems. Numer. Algo. **81**, 773–811 (2019)
23. Hansen, P.C., Jørgensen, J.S.: AIR Tools II: Algebraic Iterative Reconstruction Methods. Improved Implementation. Numerical Algorithms (2018). https://doi.org/10.1007/s11075-017-0430-x. Sotfware available at https://github.com/jakobsj/AIRToolsII/
24. Donatelli, M., Reichel, L.: Square smoothing regularization matrices with accurate boundary conditions. J. Comput. Appl. Math. **272**, 334–349 (2014)

Authors and Affiliations

Alessandro Buccini[1] · Silvia Gazzola[2] · Lucas Onisk[3] · Mirjeta Pasha[4] · Lothar Reichel[5]

✉ Alessandro Buccini
alessandro.buccini@unica.it

Silvia Gazzola
silvia.gazzola@unipi.it

Lucas Onisk
lonisk@emory.edu

Mirjeta Pasha
mpasha@vt.edu

Lothar Reichel
reichel@math.kent.edu

1 Department of Mathematics and Computer Science, University of Cagliari, Cagliari 09124, Italy

2 Department of Mathematics, University of Pisa, Pisa 56127, Italy

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

4 Department of Mathematics, Virginia Polytechnic and State University, Blacksburg 24061, VA, USA

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