Theoretical and numerical aspects of a non-stationary preconditioned iterative method for linear discrete ill-posed problems

Alessandro Buccini[a], Marco Donatelli[b], Lothar Reichel[c]

[a]*Department of Mathematics and Computer Science, University of Cagliari, 09124 Cagliari, Italy*
[b]*Department of Science and High Technology, University of Insubria, 22100 Como, Italy*
[c]*Department of Mathematical Sciences, Kent State University, Kent, OH 44242, USA*

Abstract

This work considers some theoretical and computational aspects of the recent paper [Buccini, A., Donatelli, M., Reichel, L., & Zhang, W. H. (2021), *A new nonstationary preconditioned iterative method for linear discrete ill-posed problems with application to image deblurring*, Numerical Linear Algebra with Applications, 28(2), e2353], whose aim was to relax the convergence conditions in a previous work by Donatelli and Hanke, and thereby make the iterative method discussed in the latter work applicable to a larger class of problems. This aim was achieved in the sense that the iterative method presented convergences for a larger class of problems. However, while the analysis presented is correct, it does not establish the superior behavior of the iterative method described. The present note describes a slight modification of the analysis that establishes the superiority of the iterative method. The new analysis allows to discuss the behavior of the algorithm when varying the involved parameters, which is also useful for their empirical estimation.

Keywords: discrete ill-posed inverse problems, iterated Tikhonov method, spectral equivalence
2010 MSC: 65F10, 65F22, 65F08

Email addresses: `alessandro.buccini@unica.it` (Alessandro Buccini), `marco.donatelli@uninsubria.it` (Marco Donatelli), `reichel@math.kent.edu` (Lothar Reichel)

1. Introduction

We would like to compute the solution of the least-squares problem

$$\min_{\mathbf{x}} \|A\mathbf{x} - \mathbf{b}\|^2, \tag{1}$$

where the singular values of the matrix $A \in \mathbb{R}^{m \times n}$ decay to zero without a significant gap. This makes A severely ill-conditioned. In particular, A may be rank-deficient. The vector $\mathbf{b} \in \mathbb{R}^m$ contains error-free data, and the solution $\mathbf{x} \in \mathbb{R}^n$ represents the unknown signal that we would like to determine. Throughout this paper, $\|\cdot\|$ denotes the Euclidean norm. If the solution of (1) is not unique, then we choose the solution of minimal minimal Euclidean norm. We can express this solution as $\mathbf{x}_{\text{true}} = A^\dagger \mathbf{b}$, where $A^\dagger$ denotes the Moore-Penrose pseudoinverse of A. Problems of the form (1) are commonly referred to as discrete ill-posed problems; see, e.g., [1]. They arise, for instance, from the discretization of ill-posed problems, such as Fredholm integral equations of the first kind; see, e.g., [1, 2].

In many applications, the vector $\mathbf{b}$ is not available. Instead, an error-contaminated version $\mathbf{b}^\delta$ of $\mathbf{b}$ is known and the least-squares problem (1) is replaced by

$$\min_{\mathbf{x}} \|A\mathbf{x} - \mathbf{b}^\delta\|^2. \tag{2}$$

However, due to the ill-conditioning of A, the solution of (2) of minimal Euclidean norm, namely $A^\dagger \mathbf{b}^\delta$, usually is a very poor approximation of the desired vector $\mathbf{x}_{\text{true}}$. This difficulty can be remedied by replacing the least-squares problem (2) by a nearby problem, whose solution is less sensitive to the error $\mathbf{e} = \mathbf{b}^\delta - \mathbf{b}$ in $\mathbf{b}^\delta$. We will sometimes refer to $\mathbf{e}$ as noise. The replacement is commonly referred to as *regularization*. In its simplest form, Tikhonov regularization replaces the problem (2) by the penalized least-squares problem

$$\min_{\mathbf{x}} \left\{ \|A\mathbf{x} - \mathbf{b}^\delta\|^2 + \mu \|\mathbf{x}\|^2 \right\}, \tag{3}$$

where the *regularization parameter* $\mu \geq 0$ determines how sensitive the solution $\mathbf{x}_\mu$ of (3) is to the error $\mathbf{e}$ in $\mathbf{b}^\delta$, and how close $\mathbf{x}_\mu$ is to $\mathbf{x}_{\text{true}}$. Generally, a suitable value of μ is not known a priori, but is determined during the solution of (3).

To avoid having to know a suitable value of μ before solving (3), we can solve the Tikhonov problem (3) repeatedly starting from an initial approximation of $\mathbf{x}_{\text{true}}$ and defining a sequence $\{\mu_k\}$ of regularization parameters.

The resulting *iterated Tikhonov method* is given by

$$\mathbf{x}_{k+1} = \mathbf{x}_k + A^T(AA^T + \mu_k I)^{-1}(\mathbf{b}^\delta - A\mathbf{x}_k), \qquad k = 0, 1, 2, \dots, \qquad (4)$$

where the parameters $\mu_k > 0$ are user-defined and may be chosen according to the amount of noise that corrupts the data vector $\mathbf{b}^\delta$, and $\mathbf{x}_0$ is a suitable initial vector, e.g., $\mathbf{x}_0 = \mathbf{b}^\delta$ or $\mathbf{x}_0 = A^T\mathbf{b}^\delta$. If the parameters μ_k are chosen so that

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

then the iterates converge to $A^\dagger\mathbf{b}^\delta$; see [3]. Regularization is achieved by terminating the iterations (4) after a suitable (not too large) number of iterations. One of the most popular stopping criteria is the *discrepancy principle*, which prescribes that the iterations be terminated at iteration $\hat{k}$ if

$$\left\| \mathbf{b}^\delta - A\mathbf{x}_{\hat{k}} \right\| \leq \tau\delta \quad \text{and} \quad \left\| \mathbf{b}^\delta - A\mathbf{x}_k \right\| > \tau\delta \quad \forall k \leq \hat{k},$$

where $\tau > 1$ is a user-defined parameter and

$$\|\mathbf{e}\| \leq \delta, \qquad (5)$$

is a fairly accurate estimate of the norm of the noise.

The iterated Tikhonov algorithm can be seen as a preconditioned Landweber iteration, where the preconditioner is chosen to approximate a regularized version of the pseudo-inverse of A. Other possible regularizing preconditioner are described in [4, 5, 6].

Application of the iterated Tikhonov method to large-scale problems presents two difficulties that have to be addressed: The solution of linear systems of equations with the matrix $(AA^T + \mu_k I)$ at each iteration may be computationally expensive, and the parameter μ_k has to be determined at each iteration. Donatelli and Hanke [7] proposed an approach to overcome these challenges. Their *Approximated Iterated Tikhonov (AIT) method* provides a rule for choosing the parameter μ_k that ensures the regularization properties of the method, and the matrix $A^T(AA^T + \mu_k I)^{-1}$ in (4) is replaced by $C^T(CC^T + \mu_k I)^{-1}$, where C is an approximation of A chosen so that linear systems of equations with the matrix $(CC^T + \mu_k I)$ can be solved quite rapidly.

To show convergence of the AIT method, Donatelli and Hanke [7] require the matrix C to be spectrally equivalent to A; see below. However, this is a

3

strong requirement that is almost never satisfied in practice. In applications, this condition is relaxed by the presence of noise and by a good choice of the initial guess, such that the parameter ρ, which measures the accuracy of the spectral approximation (and is used in Assumptions 1, 2, and 3 below), can be chosen small enough to yield accurate reconstructions. To reduce the sensitivity of the AIT method to the choice of ρ, i.e., to the parameter that measures how accurately C approximates A, a *modified AIT (MAIT) method*, with different theoretical requirements was proposed in [8]. In particular, a new parameter β was introduced, so that ρ can be simply set small enough, e.g., $\rho = 10^{-3}$, without compromising the numerical performances of the algorithm. Computed examples in [8] show the MAIT method to be less sensitive than the AIT method in [7] to the choice of ρ and to the noise in the available data $\mathbf{b}^\delta$. In particular, AIT may fail to converge when the amount of noise in $\mathbf{b}^\delta$ is small, while MAIT converges for any noise level, provided that the parameter β is tuned appropriately. However, as we will show in this paper, the requirements imposed in the convergence analysis of the MAIT method turn out to be equivalent to the requirement of spectral equivalence in [7].

This paper first shows that the hypotheses used for the convergence analysis in [7] and [8] are equivalent. Subsequently, we propose a weaker hypothesis and prove that the MAIT method has the same convergence properties under the new assumption as under the assumption in [8]. It was observed in [7, 9] that the choice of the initial approximate solution is of vital importance for the computation of an accurate solution, however, no rigorous mathematical reason was given. The new weaker hypothesis of the present paper provides a theoretical explanation for the need of a good initial guess. This is of great relevance in applications of the proposed method as it furnishes a practical guide for a user who applies the method. Moreover, we propose an empirical approach to estimate the parameter β. The resulting MAIT method is robust and does not require a user to explicitly determine any parameters; see numerical experiments reported in Section 4. The capability of automatically determining the parameter β is important because this allows the method to be used by non-experts in the field.

This paper is structured as follows: Section 2 briefly describes the AIT and MAIT methods. In Section 3 we introduce our new hypothesis on the relationship between the matrices A and C, and we show the same properties of the MAIT method as those shown in [8] under stronger assumptions. Section 4 discusses the choice of the initial guess for the MAIT method and

4

of the parameter β used in the algorithm, and presents a few computed examples. Finally, we draw some conclusions in Section 5.

2. The AIT and MAIT methods

We briefly describe the AIT and MAIT methods from [7] and [8], respectively, starting with the former.

2.1. The AIT method

Assumption 1 (Spectral equivalence). *Let $A, C \in \mathbb{R}^{m \times n}$, and assume that for some constant $\rho \in \left(0, \frac{1}{2}\right)$ and for all $\mathbf{z} \in \mathbb{R}^n$ it holds that*

$$\|(A - C)\mathbf{z}\| \leq \rho \|A\mathbf{z}\| .$$

Under Assumption 1 we can formulate the AIT method described by Algorithm 1.

Algorithm 1: AIT

---

1 Let A and C satisfy Assumption 1 for a given $0 < \rho < 1/2$, and fix $q \in (2\rho, 1)$. Let $\delta > 0$ satisfy (5) and let $\mathbf{x}_0$ be an initial approximation of $\mathbf{x}_{\text{true}}$;

2 $\mathbf{r}_0 = \mathbf{b}^\delta - A\mathbf{x}_0$;

3 $\tau = \frac{1+2\rho}{1-2\rho}$;

4 **for** $k = 0, 1, \ldots$ **do**

5     $\tau_k = \|\mathbf{r}_k\| / \delta$;

6     $q_k = \max\{q, 2\rho + (1 + \rho)/\tau_k\}$;

7   Determine μ_k such that $\left\|\mathbf{r}_k - CC^T(CC^T + \mu_k I)^{-1}\mathbf{r}_k\right\| = q_k \|\mathbf{r}_k\|$;

8     $\mathbf{h}_k = C^T(CC^T + \mu_k I)^{-1}\mathbf{r}_k$;

9     $\mathbf{x}_{k+1} = \mathbf{x}_k + \mathbf{h}_k$;

10     $\mathbf{r}_{k+1} = \mathbf{b}^\delta - A\mathbf{x}_{k+1}$;

11   **if** $\|\mathbf{r}_{k+1}\| \leq \tau\delta$ **then**

12   exit;

Throughout this paper,

$$\mathbf{e}_k = \mathbf{x}_{\text{true}} - \mathbf{x}_k$$

denotes the error in the kth iterate $\mathbf{x}_k$ generated by the algorithms for the AIT or MAIT methods.

Proposition 1 ([7]). *Let Assumption 1 hold. Then the norm of the error $\mathbf{e}_k$ in the iterate $\mathbf{x}_k$ generated by Algorithm 1 decreases monotonically, namely*

$$\|\mathbf{e}_k\|^2 - \|\mathbf{e}_{k+1}\|^2 \geq 2\rho\|(CC^T + \mu_k I)^{-1}\mathbf{r}_k\| \, \|\mathbf{r}_k\|,$$

as long as $\|\mathbf{r}_k\| > \tau\delta$, where $\mathbf{r}_k = \mathbf{b}^\delta - A\mathbf{x}_k$.

Corollary 2 ([7]). *Under the assumptions of Proposition 1, let k_δ denote index of the last iterate determined by Algorithm 1. Then*

$$\|\mathbf{e}_0\|^2 \geq 2\rho \sum_{k=0}^{k_\delta-1} \|(CC^T + \mu_k I)^{-1}\mathbf{r}_k\| \, \|\mathbf{r}_k\| \geq c \sum_{k=0}^{k_\delta-1} \|\mathbf{r}_k\|^2,$$

for some constant $c > 0$ that only depends on ρ and q, where q is defined in Algorithm 1.

The above corollary implies that if $\delta > 0$, then the computations with Algorithm 1 terminate after a finite number of iterations. The following theorem is concerned with the solution of the noise-free problem (1) and shows that when $\delta = 0$, the iterates generated by Algorithm 1 converge to a solution of the exact problem (with data vector $\mathbf{b}$) that is closest to $\mathbf{x}_0$. In particular, if $\mathbf{x}_0 = \mathbf{0}$ and $\delta = 0$, then the iterates of Algorithm 1 converge to $\mathbf{x}_{\text{true}}$.

Theorem 3 ([7]). *Assume that $\mathbf{x}_0$ is not a solution of (1). Then the sequence of iterates $\mathbf{x}_k$, $k = 0, 1, 2, \ldots$, generated by Algorithm 1 with $\mathbf{b}^\delta$ replaced by $\mathbf{b}$ converges as $k \to \infty$ to a solution of (2) that is closest in Euclidean norm to $\mathbf{x}_0$.*

It follows from the next theorem that the AIT algorithm is an iterative regularization method.

Theorem 4 ([7]). *Let Assumption 1 be valid and let $\delta \mapsto \mathbf{b}^\delta$ be a function such that (5) holds for all $\delta > 0$. For fixed parameters τ and q (defined in Algorithm 1), let $\mathbf{x}^\delta$ denote the approximate solution computed by Algorithm 1. Then, choosing $\mathbf{x}_0 = \mathbf{0}$, as $\delta \to 0$, $\mathbf{x}^\delta$ converges to $\mathbf{x}_{\text{true}} = A^\dagger \mathbf{b}$.*

2.2. The MAIT method

We now describe the MAIT method proposed in [8], where the following condition is imposed on the matrices A and C.

Assumption 2. *Let $A, C \in \mathbb{R}^{m \times n}$ and assume that there are constants $\rho \in \left(0, \frac{1}{2}\right)$ and $\beta > 0$, such that for all $\mathbf{z} \in \mathbb{R}^n$ it holds that*

$$\|(A - C)\mathbf{z}\| \leq \rho\left(\|A\mathbf{z}\| + \beta\right).$$

The above assumption is used in [8] to prove the convergence of the MAIT method. We next show that Assumptions 1 and 2 are equivalent.

Proposition 5. *Assumptions 1 and 2 are equivalent.*

Proof. Obviously, if the matrices A and C satisfy Assumption 1, then they satisfy Assumption 2 for any $\beta > 0$. Assume instead that the matrices A and C satisfy Assumption 2 for a finite $\beta > 0$, but that they do not satisfy Assumption 1. Then there is a vector $\bar{\mathbf{z}} \in \mathbb{R}^n$ such that

$$\|(A - C)\bar{\mathbf{z}}\| > \rho\|A\bar{\mathbf{z}}\|, \tag{6}$$

for all $\rho \in \left(0, \frac{1}{2}\right)$. However, by Assumption 2, for all $\lambda > 0$, we have that

$$\|(A - C)\lambda\bar{\mathbf{z}}\| \leq \rho\left(\|A\lambda\bar{\mathbf{z}}\| + \beta\right).$$

Therefore, it holds for all $\lambda > 0$ that

$$\beta \geq \frac{\lambda}{\rho}\left(\|(A - C)\bar{\mathbf{z}}\| - \rho\|A\bar{\mathbf{z}}\|\right).$$

Letting $\lambda \to \infty$, it follows from (6) that $\beta > 0$ cannot be finite. This contradiction shows that Assumption 2 implies Assumption 1. $\square$

Computed examples reported in [8] illustrate that the parameter β is important as it allows the iterates generated by the MAIT method to converge when the noise level $\delta > 0$ is very small, in situations when the iterates determined by the AIT methods do not. This indicates that the MAIT method may not require Assumption 2 to hold to yield convergence. A new weaker assumption that secures convergence of the MAIT method is derived in Section 3. We would like to stress that, even though Assumption 2 is not weaker than Assumption 1 as stated in [8], the theoretical results in [8] are still valid.

Theorem 6 ([8]). *Let Assumption 2 hold. With the notation of Algorithm 2, we have that, if*

$$\|\mathbf{r}_{k+1}\| > \tau\delta \quad \text{and} \quad \|\mathbf{r}_{k+1}\| > \tau\frac{\delta + \beta}{t_0},$$

Algorithm 2: MAIT

---

1 Let the matrices A and C satisfy Assumption 2 for some $0 < \rho < 1/2$ and $\beta \geq 0$. Fix $q \in (2\rho, 1)$, let $\delta > 0$ satisfy (5), and let $\mathbf{x}_0$ be an initial approximation of $\mathbf{x}_{\text{true}}$;

2 $t_0 = \min\left\{\frac{\delta}{\beta}, \frac{\beta}{\delta}\right\} + 1$;

3 $\mathbf{r}_0 = \mathbf{b}^\delta - A\mathbf{x}_0$;

4 $\tau = \frac{1+2\rho}{1-2\rho}$;

5 for $k = 0, 1, \dots$ **do**

6 $\quad$ $\tau_k = t_0 \dfrac{\|\mathbf{r}_k\|}{\delta + \beta}$ $\quad q_k = \max\{q, 2\rho + (1+\rho)/\tau_k\}$;

7 $\quad$ Determine μ_k such that $\|\mathbf{r}_k - CC^T(CC^T + \mu_k I)^{-1}\mathbf{r}_k\| = q_k\|\mathbf{r}_k\|$

$\quad$ $\mathbf{h}_k = C^T(CC^T + \mu_k I)^{-1}\mathbf{r}_k$;

8 $\quad$ $\mathbf{x}_{k+1} = \mathbf{x}_k + \mathbf{h}_k$;

9 $\quad$ $\mathbf{r}_{k+1} = \mathbf{b}^\delta - A\mathbf{x}_{k+1}$;

10 $\quad$ **if** $\|\mathbf{r}_{k+1}\| \leq \tau\delta$ *or* $\|\mathbf{r}_{k+1}\| \leq \tau\dfrac{\delta + \beta}{t_0}$ **then**

11 $\quad\quad$ exit;

---

then

$$\|\mathbf{e}_k\|^2 - \|\mathbf{e}_{k+1}\|^2 \geq 2\rho \frac{\tau - 1}{\tau} \|(CC^T + \mu_k I)^{-1}\mathbf{r}_k\| \, \|\mathbf{r}_k\|.$$

In particular, it follows from $\tau > 1$, that $\|\mathbf{e}_k\|$ decreases monotonically as k increases.

3. A new convergence proof

We first introduce a new relaxed condition for the convergence of the MAIT method.

Assumption 3. *Let the matrices $A, C \in \mathbb{R}^{m \times n}$. Assume that there are constants $\rho \in \left(0, \frac{1}{2}\right)$, $\beta > 0$, and $\gamma > 0$ such that for all $\mathbf{z} \in \mathbb{R}^n$ with $\|\mathbf{z}\| \leq \gamma$ it holds*

$$\|(A - C)\mathbf{z}\| \leq \rho \left(\|A\mathbf{z}\| + \beta\right).$$

Assumption 3 is obtained by requiring that the condition in Assumption 2 only holds for vectors $\mathbf{z} \in \mathbb{R}^n$ in a ball centered at $\mathbf{0}$ with radius γ. The new parameter γ is not directly involved in the algorithm.

We first observe that [8, Lemma 1], which was used to justify the stopping criterion in Algorithm 2, still holds, since its proof does not use Assumption 2. Therefore, we do not dwell on the stopping criterion here.

We now show an auxiliary result that will be needed in the following. This result is the analogue of Proposition 3 in [8].

Proposition 7. *Assume that Assumption 3 holds and that $\|\mathbf{e}_k\| \leq \gamma$, where γ is defined in Assumption 3. Let $\tau_k = t_0 \frac{\|\mathbf{r}_k\|}{\delta + \beta}$, where $t_0 = \min\left\{\frac{\beta}{\delta}, \frac{\delta}{\beta}\right\} + 1$ is defined in Algorithm 2. Then*

$$\|\mathbf{r}_k - C\mathbf{e}_k\| \leq \left(\rho + t_0 \frac{1 + \rho}{\tau_k}\right) \|\mathbf{r}_k\| - \beta.$$

Proof. The proof is almost identical to the proof of [8, Proposition 3]. We

9

report it here for completeness. We have

$$\begin{aligned}
\|\mathbf{r}_k - C\mathbf{e}_k\| &= \|\mathbf{b}^\delta - \mathbf{b} + \mathbf{b} - A\mathbf{x}_k - C\mathbf{e}_k\| \\
&= \|\mathbf{b}^\delta - \mathbf{b} + A(\mathbf{x}_{\text{true}} - \mathbf{x}_k) - C\mathbf{e}_k\| \\
&\leq \|\mathbf{b}^\delta - \mathbf{b}\| + \|(A - C)(\mathbf{x}_{\text{true}} - \mathbf{x}_k)\| \\
&\stackrel{(a)}{\leq} \|\mathbf{b}^\delta - \mathbf{b}\| + \rho(\|A(\mathbf{x}_{\text{true}} - \mathbf{x}_k)\| + \beta) \\
&\leq \|\mathbf{b}^\delta - \mathbf{b}\| + \rho(\|\mathbf{b} - \mathbf{b}^\delta + \mathbf{b}^\delta - A\mathbf{x}_k\| + \beta) \\
&\leq (1 + \rho)\delta + \rho\|\mathbf{r}_k\| + \rho\beta,
\end{aligned} \tag{7}$$

where inequality (a) follows from Assumption 3, since we assumed that $\|\mathbf{e}_k\| \leq \gamma$.

It follows from the definition of τ_k that

$$\delta = t_0 \frac{\|\mathbf{r}_k\|}{\tau_k} - \beta.$$

Substituting this expression into (7), we obtain

$$\|\mathbf{r}_k - C\mathbf{e}_k\| \leq (1+\rho)\left(t_0\frac{\|\mathbf{r}_k\|}{\tau_k} - \beta\right) + \rho\|\mathbf{r}_k\| + \rho\beta = \left(\rho + t_0\frac{1+\rho}{\tau_k}\right)\|\mathbf{r}_k\| - \beta.$$

$$\square$$

We can now show our main theoretical result.

Theorem 8. *Let Assumption 3 hold and assume that $\|\mathbf{e}_k\| \leq \gamma$. With the notation of Algorithm 2, we have that, if*

$$\|\mathbf{r}_{k+1}\| > \tau\delta \quad and \quad \|\mathbf{r}_{k+1}\| > \tau\frac{\delta + \beta}{t_0},$$

then

$$\|\mathbf{e}_k\|^2 - \|\mathbf{e}_{k+1}\|^2 \geq 2\rho\frac{\tau - 1}{\tau}\|(CC^T + \mu_k I)^{-1}\mathbf{r}_k\| \|\mathbf{r}_k\|.$$

It follows from $\tau > 1$, that $\|\mathbf{e}_k\|$ decreases monotonically as k increases. In particular, since $\|\mathbf{e}_k\| \leq \gamma$, we also have $\|\mathbf{e}_{k+1}\| \leq \gamma$.

Proof. The proof of this result is almost identical to the proof of [8, Theorem 3]. Let $\langle \mathbf{v}, \mathbf{u} \rangle$ denote the standard inner product in $\mathbb{R}^n$ and let

$$\mathbf{h}_k = C^T(CC^T + \mu_k I)^{-1}\mathbf{r}_k.$$

10

Then

$$\|\mathbf{e}_k\|^2 - \|\mathbf{e}_{k+1}\|^2 = 2\langle \mathbf{e}_k, \mathbf{h}_k \rangle - \|\mathbf{h}_k\|^2$$
$$\geq 2\langle \mathbf{e}_k, C^T(CC^T + \mu_k I)^{-1}\mathbf{r}_k \rangle - 2\langle \mathbf{r}_k, CC^T(CC^T + \mu_k I)^{-2}\mathbf{r}_k \rangle$$
$$= 2\langle \mathbf{r}_k, (CC^T + \mu_k I)^{-1}\mathbf{r}_k \rangle - 2\langle \mathbf{r}_k - C\mathbf{e}_k, (CC^T + \mu_k I)^{-1}\mathbf{r}_k \rangle$$
$$\quad - 2\langle \mathbf{r}_k, CC^T(CC^T + \mu_k I)^{-2}\mathbf{r}_k \rangle$$
$$= 2\langle \mathbf{r}_k, \mu_k(CC^T + \mu_k I)^{-2}\mathbf{r}_k \rangle - 2\langle \mathbf{r}_k - C\mathbf{e}_k, (CC^T + \mu_k I)^{-1}\mathbf{r}_k \rangle$$
$$\geq 2\|(CC^T + \mu_k I)^{-1}\| \left(\|\mu_k(CC^T + \mu_k I)^{-1}\| - \|\mathbf{r}_k - C\mathbf{e}_k\| \right).$$

By the definition of $\mathbf{h}_k$, we have

$$\mathbf{r}_k - C\mathbf{h}_k = \mathbf{r}_k - CC^T(CC^T + \mu_k I)^{-1}\mathbf{r}_k = \mu_k(CC^T + \mu_k I)^{-1}\mathbf{r}_k.$$

Combining the above inequality and equality, and using the definitions of μ_k and τ_k, as well as Proposition 7 (which can be applied since we assumed that $\|\mathbf{e}_k\| \leq \gamma$), we obtain

$$\|\mathbf{e}_k\|^2 - \|\mathbf{e}_{k+1}\|^2 \geq 2\|(CC^T + \mu_k I)^{-1}\| \left(\|\mathbf{r}_k - C\mathbf{h}_k\| - \|\mathbf{r}_k - C\mathbf{e}_k\| \right)$$
$$\geq 2\|(CC^T + \mu_k I)^{-1}\| \left(q_k\|\mathbf{r}_k\| - ((1+\rho)\delta + \rho\|\mathbf{r}_k\| + \rho\beta) \right)$$
$$\geq 2\|(CC^T + \mu_k I)^{-1}\| \left(\left(2\rho + \frac{1+\rho}{\tau_k} - \rho\right)\|\mathbf{r}_k\| - (1+\rho)\delta - \rho\beta \right)$$
$$\geq 2\|(CC^T + \mu_k I)^{-1}\| \left(\rho\|\mathbf{r}_k\| + \frac{1+\rho}{t_0\|\mathbf{r}_k\|}(\delta + \beta)\|\mathbf{r}_k\| - (1+\rho)\delta - \rho\beta \right)$$
$$= 2\|(CC^T + \mu_k I)^{-1}\| \left(\rho\|\mathbf{r}_k\| + \frac{1+\rho}{t_0}(\delta + \beta) - (1+\rho)\delta - \rho\beta \right).$$

We now show that

$$\frac{1+\rho}{t_0}(\delta + \beta) - (1+\rho)\delta - \rho\beta \geq -\rho\delta.$$

First consider the expression $(\delta + \beta)/t_0$. If $\delta \geq \beta$, then $\delta/\beta \geq \beta/\delta$ and, thus,

$$\frac{\delta + \beta}{t_0} = \frac{\delta + \beta}{\beta/\delta + 1} = \delta. \tag{8}$$

Conversely, if $\delta \leq \beta$, then $\delta/\beta \leq \beta/\delta$ and, therefore,

$$\frac{\delta + \beta}{t_0} = \frac{\delta + \beta}{\delta/\beta + 1} = \beta. \tag{9}$$

Combining the expressions (8) and (9), we obtain

$$\frac{\delta + \beta}{t_0} = \max\{\beta, \delta\}.$$

We now can bound

$$\frac{1 + \rho}{t_0}(\delta + \beta) - (1 + \rho)\delta - \rho\beta$$

below. Thus,

$$\frac{1 + \rho}{t_0}(\delta + \beta) - (1 + \rho)\delta - \rho\beta = (1 + \rho)\max\{\beta, \delta\} - (1 + \rho)\delta - \rho\beta$$

$$= (\max\{\beta, \delta\} - (1 + \rho)\delta) + (\rho\max\{\beta, \delta\} - \rho\beta)$$

$$\geq (\delta - (1 + \rho)\delta) + (\rho\beta - \rho\beta) = -\rho\delta.$$

Recalling that, by assumption, $\|\mathbf{r}_k\| \geq \tau\delta$, we obtain

$$\|\mathbf{e}_k\|^2 - \|\mathbf{e}_{k+1}\|^2 \geq 2\|(CC^T + \mu_k I)^{-1}\|\,(\rho\|\mathbf{r}_k\| - \rho\delta)$$

$$\geq 2\|(CC^T + \mu_k I)^{-1}\|\left(\rho\|\mathbf{r}_k\| - \rho\frac{\|\mathbf{r}_k\|}{\tau}\right)$$

$$= 2\|(CC^T + \mu_k I)^{-1}\|\rho\frac{\tau - 1}{\tau}\|\mathbf{r}_k\|.$$

$\square$

The following result follows trivially.

Corollary 9. *With the notation of Theorem 8, if $\|\mathbf{e}_0\| \leq \gamma$ then, for all $k \geq 0$ such that*

$$\|\mathbf{r}_{k+1}\| > \tau\delta \quad and \quad \|\mathbf{r}_{k+1}\| > \tau\frac{\delta + \beta}{t_0},$$

it holds

$$\|\mathbf{e}_{k+1}\| \leq \|\mathbf{e}_k\|.$$

Remark 1. Note that Corollary 9 requires that

$$\|\mathbf{e}_0\| \leq \gamma,$$

i.e., that the initial approximate solution $\mathbf{x}_0$ is not too far from $\mathbf{x}_{\text{true}}$. This illustrates why in [7] the choice of the initial approximate solution is important and the authors avoided using $\mathbf{x}_0 = \mathbf{0}$ in the numerical examples. This was pointed out in [9] as well, where the author selected $A^T\mathbf{b}^\delta$ as the initial approximate solution.

Remark 2. It is stated in [8] that the MAIT method is *not* an iterative regularization method under Assumption 2, i.e., if we let $\mathbf{x}^\delta$ denote the computed approximation determined by the method when the error in the data $\mathbf{b}^\delta$ is bounded by δ, then, in general, it does not hold that $\mathbf{x}^\delta \to A^\dagger \mathbf{b}$ as $\delta \to 0$. This depends on the parameter β. Obviously, the MAIT method is not an iterative regularization method under Assumption 3 either, since the latter assumption is weaker.

4. Numerical results

This section presents some numerical results. First, we consider an example in one space-dimension, where the task is to solve a linear system of equations that is obtained by discretizing a Fredholm integral equation of the first kind of the form

$$g(s) = \int_a^b k(s - t)x(t)dt; \tag{10}$$

see below. We will show the behavior of both the AIT and MAIT methods when applied to problems in one and two space-dimensions as $\delta \to 0$, highlighting that the MAIT method, as already illustrated in [8], converges also for small values of δ, while the AIT method may fail to do so.

It follows from Proposition 7 that, if Assumption 3 is satisfied, then, for all k, it holds that

$$\frac{\|\mathbf{r}_k - C\mathbf{e}_k\| + \beta}{\|\mathbf{r}_k\|} \le \rho + t_0 \frac{1 + \rho}{\tau_k},$$

provided that $\|\mathbf{e}_0\| \le \gamma$. The left-hand side of this inequality depends on β. We will consider the simplified version

$$\frac{\|\mathbf{r}_k - C\mathbf{e}_k\|}{\|\mathbf{r}_k\|} \le \rho + t_0 \frac{1 + \rho}{\tau_k}, \tag{11}$$

which is closely related to the inequality used in [7, Section 5]. In this section we illustrate that inequality (11) holds with small values of $\rho > 0$, and how the choices of $\mathbf{x}_0$ and β affect the inequality. In particular, for a fixed ρ, we elucidate the following:

- Inequality (11) is not satisfied for an imprudent choice of $\mathbf{x}_0$.

Table 1: Description of the numerical examples.

True image	Size of the true image	PSF	Size of the PSF
Phantom	237×237 pixels	Gaussian	237×237 pixels
Cameraman	248×248 pixels	Average	18×18 pixels
Clock	248×248 pixels	Motion	7×7 pixels
Grain	238×238 pixels	Hand-shaking	17×17 pixels

- We can estimate the value of β in Assumption 3 by exploiting (11).

To this aim, we analyze the same computed examples as in [8]. In particular, we consider the four image deblurring problems described in Table 1 and use two noise levels, namely 0.1% and 1%, where we say that the data $\mathbf{b}^\delta$ is contaminated by noise $\mathbf{e}$ of level σ if

$$\sigma = \frac{\delta}{\|\mathbf{b}\|}.$$

In the computed examples, the noise $\mathbf{e}$ in $\mathbf{b}^\delta$ is white Gaussian.

The point spread functions (PSFs) model four different kinds of blur descibed in Table 1. These blurs are further illustrated in [8]. We measure the relative restoration error (RRE) in the computed solution $\mathbf{x}$ for all examples. It is given by

$$\mathrm{RRE}(\mathbf{x}) = \frac{\|\mathbf{x} - \mathbf{x}_{\mathrm{true}}\|}{\|\mathbf{x}_{\mathrm{true}}\|}.$$

4.1. Example in one space-dimension

We consider a convolution problem of the form (10), where

$$k(u) = \mathrm{e}^{-u^2},$$

$a = -\pi$, and $b = \pi$. By discretizing the problem with a collocation method with equidistant collocation points, we obtain a Toeplitz matrix A, and choose C to be the circulant matrix that is generated by the same symbol as A; see, e.g., [10], as well as [11, 12], for related results. We use a grid of 3001 equispaced points and obtain $A \in \mathbb{R}^{3001 \times 3001}$. The vector $\mathbf{x}_{\mathrm{true}}$ is constructed as a uniform sampling of $f(x) = \sin(x)$ in $[-\pi, \pi]$ and obtain the

Table 2: RRE values obtained for different noise levels with AIT and MAIT for the example in one space-dimension.

Method	Noise Level	
	$\sigma = 0.1\%$	$\sigma = 1\%$
AIT	8.2487×10^3	0.0061480
MAIT	0.0075916	0.0061480

noise-free right-hand side as $\mathbf{b} = A\mathbf{x}_{\text{true}}$. Adding white Gaussian noise with $\sigma \in \{0.1\%, 1\%\}$ to $\mathbf{b}$ gives $\mathbf{b}^\delta$. We ran both AIT and MAIT, where we set $q = 0.7$, $\rho = 10^{-2}$, $\beta = 10^{-1}$, and $\mathbf{x}_0 = \mathbf{0}$. Table 2 reports the RREs obtained for the AIT and MAIT methods for two noise levels. We can observe that AIT fails to converge for $\sigma = 0.1\%$ while MAIT produces a very accurate reconstruction of $\mathbf{x}_{\text{true}}$. For $\sigma = 1\%$ the AIT and MAIT methods provide reconstructions of the same quality..

Note that the MAIT method provides a more accurate reconstruction with $\sigma = 1\%$ than with $\sigma = 0.1\%$. This may seem counter-intuitive, but it is due to the parameter β. It is a consequence that the MAIT method is not an iterative regularization method, as already pointed out in [8].

Observe that the AIT and MAIT algorithms can be used straightforwardly when A is a Toeplitz matrix by letting C be a circulant matrix that approximates A. For a non-structured matrix A, the construction of a suitable preconditioner C is not as straightforward. A possible approach for the construction of C in this situation has recently been proposed in [13].

4.2. Behavior for $\sigma \to 0$

We now turn our attention to studying the behavior of the AIT and MAIT methods as the noise level σ decreases to zero. In particular, we are interested in showing how the algorithms select the parameters μ_k and the evolution of the RRE. We consider the Grain example in Table 1. Thus, this reconstruction problem is in two space-dimensions. We fix $\mathbf{x}_0 = \mathbf{0}$ and $\beta = 150$ and run the algorithms for several values of the noise level σ in the interval $[10^{-5}, 10^{-1}]$.

Figure 1 reports the parameters computed by AIT and MAIT for three values of σ, i.e., for 10^{-5}, 10^{-3}, and 10^{-1}. We can observe that for the first two values of σ, AIT produces very small values μ_k, which results in poor approximations of the desired solution $\mathbf{x}_{\text{true}}$. On the other hand, the μ_k-values determined by MAIT decrease during the first iterations and then
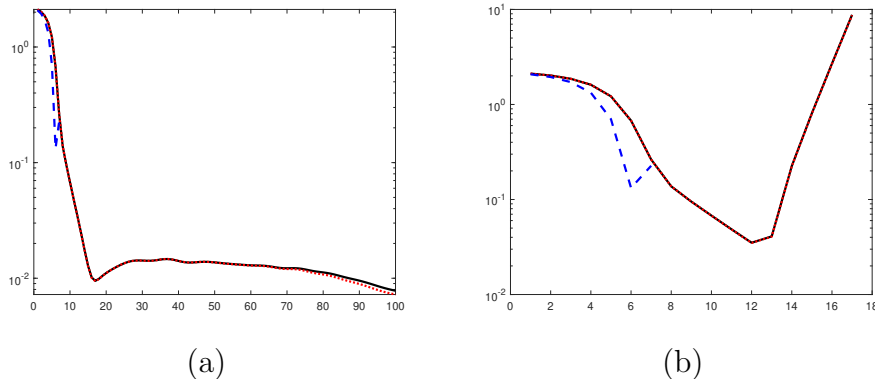
Figure 1: Values of the regularization parameter μ_k determined by AIT in panel (a) and by MAIT in panel (b) for different noise levels σ as a function of k. The solid black graph is for $\sigma = 10^{-5}$, the dotted red graph is for $\sigma = 10^{-3}$, and the dashed blue graph is for $\sigma = 10^{-1}$.

start to increase, stabilizing the computations. For $\sigma = 10^{-1}$ both methods converge quickly and the values of μ_k do not become too small.

In Figure 2 we show the evolution of the RRE versus σ for both algorithms. We can observe that for small values of σ, AIT does not produce reasonable approximations of $\mathbf{x}_{\text{true}}$ and the RRE-value is larger than 1, while MAIT produces accurate reconstructions for all σ. When σ is large enough, we observe that AIT and MAIT are equivalent (quality-wise). This was already pointed out in [8]. Finally, we note that the RRE for MAIT does not decrease to 0 as $\sigma \to 0$; this is in agreement with the fact that MAIT is not an iterative regularization method.

4.3. On the choice of the initial approximate solution $\mathbf{x}_0$

For simplicity, we fix $\rho = 10^{-3}$ and illustrate, for a fixed value of β, inequality (11) for three choices of $\mathbf{x}_0$, namely $\mathbf{0}$, $\mathbf{b}^\delta$, and $A^T\mathbf{b}^\delta$. Obviously, the choice of $\mathbf{x}_0$ affects the value of γ as discussed in Remark 1.

We first choose the fairly large value $\beta = 400$. Table 3 displays graphs with the evolution of $\frac{\|\mathbf{r}_k - C\mathbf{e}_k\|}{\|\mathbf{r}_k\|}$ and $\rho + t_0\frac{1+\rho}{\tau_k}$ as a function of the iteration number k. We can observe that for $\mathbf{x}_0 = \mathbf{0}$ the inequality (11) is not satisfied for all k, while for $\mathbf{x}_0 = \mathbf{b}^\delta$ and $\mathbf{x}_0 = A^T\mathbf{b}^\delta$ the inequality always holds. Moreover, we see that the choice of the initial approximate solution affects the number of iterations required to reach convergence, in particular, the choice $\mathbf{x}_0 = \mathbf{0}$ necessitates a much higher computational effort. This illustrates that
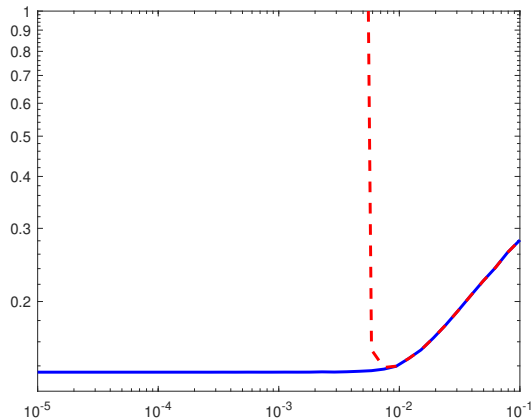
16

Figure 2: Values of the RRE obtained with AIT and MAIT for different values of the noise level σ. The dashed red graph reports results for AIT, while the solid blue graph shows results for MAIT.

the choice of the initial approximate solution is important for this method and that an imprudent choice of $\mathbf{x}_0$ may result in an computed approximation of poor quality. To quantify this, we report in Table 4 for the Cameraman example the RRE for the six different cases considered. As was already pointed out in [9], we can observe that a good choice for $\mathbf{x}_0$ is usually $A^T \mathbf{b}^\delta$. Intuitively, this is due to the fact that the operator A^T is a low-pass filter and removes some noise from the reconstruction. The results for the other examples are similar and, therefore, we do not report them.

4.4. On the choice of β

We now turn to the role of β. Let us fix $\mathbf{x}_0 = A^T \mathbf{b}^\delta$ and show how the evolution of $\frac{\|\mathbf{r}_k - C\mathbf{e}_k\|}{\|\mathbf{r}_k\|}$ and $\rho + t_0 \frac{1+\rho}{\tau_k}$ throughout the iterations changes with β. For simplicity we let $\sigma = 1\%$ and $\rho = 10^{-3}$. Our results are reported in Table 5. We can observe that for small values of β, inequality (11) does not hold, while for larger values of this parameter the inequality is satisfied. Moreover, as β increases the gap between $\frac{\|\mathbf{r}_k - C\mathbf{e}_k\|}{\|\mathbf{r}_k\|}$ and $\rho + t_0 \frac{1+\rho}{\tau_k}$ increases as well. In particular, we can see that, as β increases, the blue curve approaches the red curve from above, and for a certain value β^* the two switch places.

The tables suggest a heuristic approach to determine a suitable value of the parameter β. Let us first observe that the curves in Table 5 cannot be plotted unless the exact solution is available. Therefore, we cannot directly

17

Table 3: Verification of (11) for fixed $\beta = 400$, $\rho = 10^{-3}$, different choices of σ, and $\mathbf{x}_0$. The blue solid curves and the red dashed curves report the values of $\frac{\|\mathbf{r}_k - C\mathbf{e}_k\|}{\|\mathbf{r}_k\|}$ and $\rho + t_0 \frac{1+\rho}{\tau_k}$ versus the number of iterations, respectively. Below each graph, we report the number of iterations required to reach convergence.
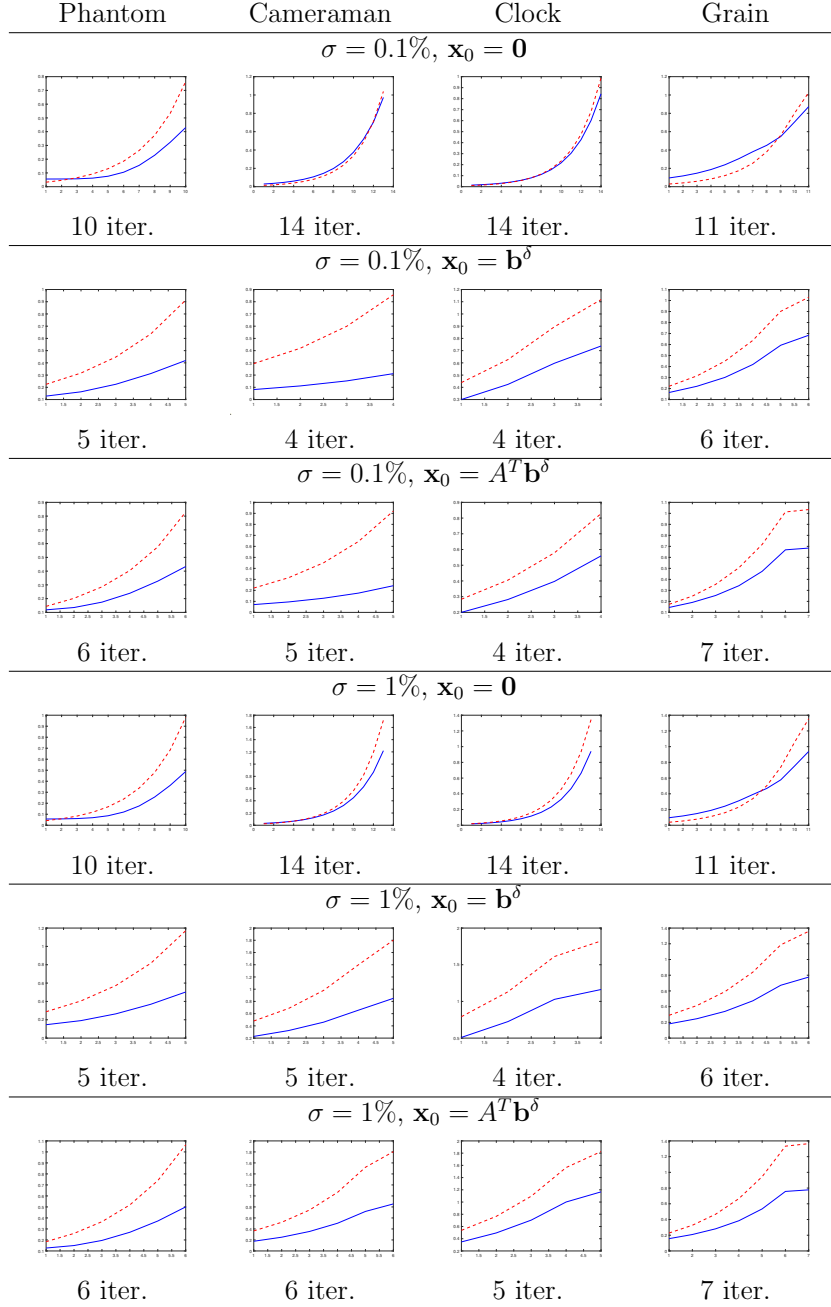
Phantom	Cameraman	Clock	Grain
$\sigma = 0.1\%$, $\mathbf{x}_0 = \mathbf{0}$			
10 iter.	14 iter.	14 iter.	11 iter.
$\sigma = 0.1\%$, $\mathbf{x}_0 = \mathbf{b}^\delta$			
5 iter.	4 iter.	4 iter.	6 iter.
$\sigma = 0.1\%$, $\mathbf{x}_0 = A^T \mathbf{b}^\delta$			
6 iter.	5 iter.	4 iter.	7 iter.
$\sigma = 1\%$, $\mathbf{x}_0 = \mathbf{0}$			
10 iter.	14 iter.	14 iter.	11 iter.
$\sigma = 1\%$, $\mathbf{x}_0 = \mathbf{b}^\delta$			
5 iter.	5 iter.	4 iter.	6 iter.
$\sigma = 1\%$, $\mathbf{x}_0 = A^T \mathbf{b}^\delta$			
6 iter.	6 iter.	5 iter.	7 iter.

18

Table 4: RREs obtained with $\beta = 400$, $\rho = 10^{-3}$, and different choices of σ and $\mathbf{x}_0$ in the Cameraman example.

Initial Guess	Noise Level	
	$\sigma = 0.1\%$	$\sigma = 1\%$
$\mathbf{x}_0 = \mathbf{0}$	0.13085	0.12184
$\mathbf{x}_0 = \mathbf{b}^\delta$	0.13124	0.11819
$\mathbf{x}_0 = A^T\mathbf{b}^\delta$	0.13056	0.11813

use them to determine β. However, one can determine a suitable value of β for synthetic examples and then use this value for the desired example. In detail, we proceed as follows:

1. Create J synthetic examples with the same matrix A, known exact solutions $\mathbf{x}_{\text{true}}^{(j)}$, and data $\mathbf{b}^{(j)}$ such that

$$\frac{\left\|A\mathbf{x}_{\text{true}}^{(j)} - \mathbf{b}^{(j)}\right\|}{\left\|A\mathbf{x}_{\text{true}}^{(j)}\right\|} = \sigma, \quad j = 1, 2, \ldots, J,$$

where σ is the noise level of the problem we would like to solve. Moreover, we assume that the solutions $\mathbf{x}_{\text{true}}^{(j)}$ are scaled similarly as the unknown solution $\mathbf{x}_{\text{true}}$. For instance, they all represent 8-bit grayscale images.

2. For each synthetic problem, we run the MAIT algorithm with several values of β and determine the smallest value of them for which inequality (11) holds. We denote these values by β_j, $j = 1, 2, \ldots, J$.

3. Use the average value

$$\hat{\beta} = \frac{1}{J} \sum_{j=1}^{J} \beta_j$$

for the real problem we would like to solve.

We will illustrate this approach when applied to the following image deblurring problem: Consider the exact image in Figure 3(a), blur it with the average PSF in Figure 3(b), and cut the borders to simulate realistic boundary conditions. This gives the blurred and noise-free image shown in Figure 3(c). We impose reflexive boundary conditions. Two noise levels are considered, namely $\sigma = 0.1\%$ and $\sigma = 1\%$. For each noise level, we generate

19

Table 5: Verification of (11) for fixed $\sigma = 1\%$, $\rho = 10^{-3}$, and $\mathbf{x}_0 = A^T\mathbf{b}^\delta$, and different choices of β. The blue solid curves and the red dashed curves report the values of $\frac{\|\mathbf{r}_k - C\mathbf{e}_k\|}{\|\mathbf{r}_k\|}$ and $\rho + t_0\frac{1+\rho}{\tau_k}$, respectively, versus the number of iterations.
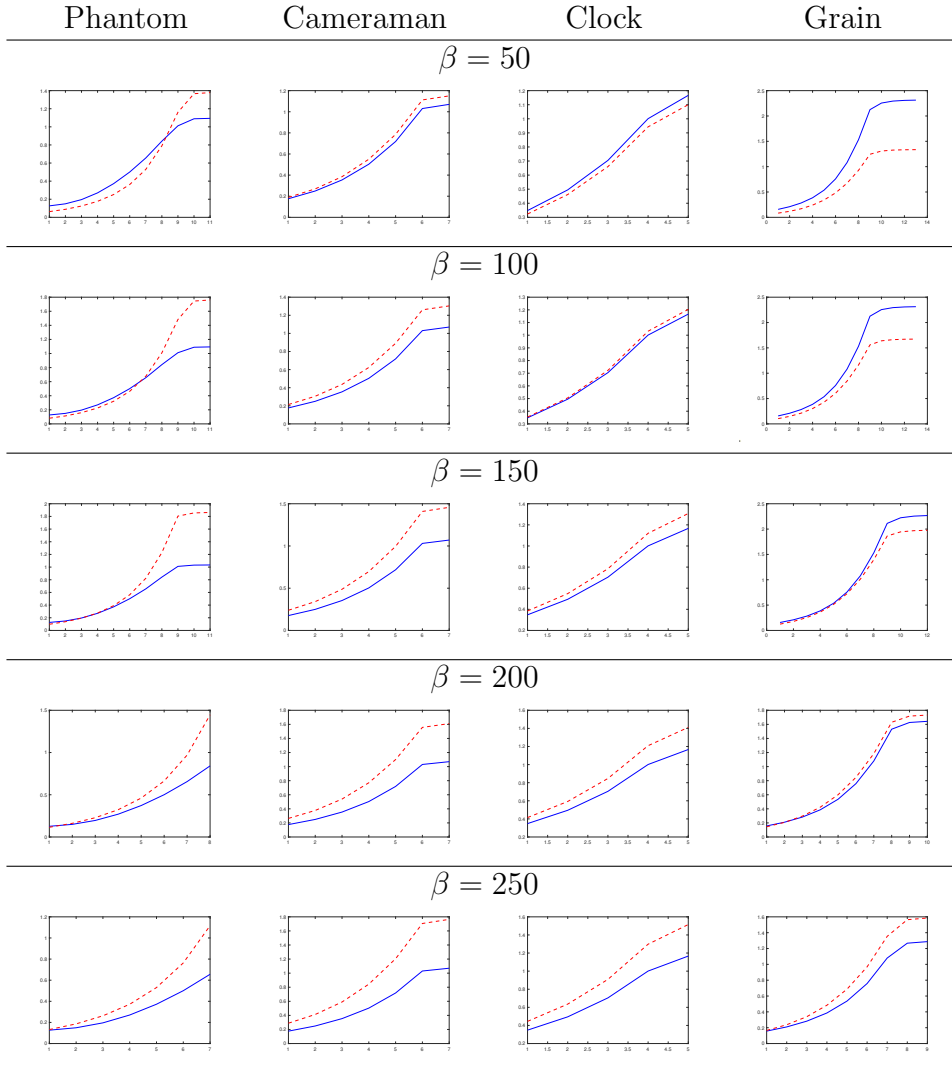
Phantom	Cameraman	Clock	Grain
$\beta = 50$			

| $\beta = 100$ | | | |

| $\beta = 150$ | | | |

| $\beta = 200$ | | | |

| $\beta = 250$ | | | |

Table 6: House test case: RRE obtained with both AIT and MAIT and the value $\hat{\beta}$ determined by our heuristic for the two considered noise levels.

Noise level	RRE AIT	RRE MAIT	$\hat{\beta}$
$\sigma = 0.1\%$	–	0.059279	187.5
$\sigma = 1\%$	0.061720	0.061720	112.5

(a)   (b)   (c)
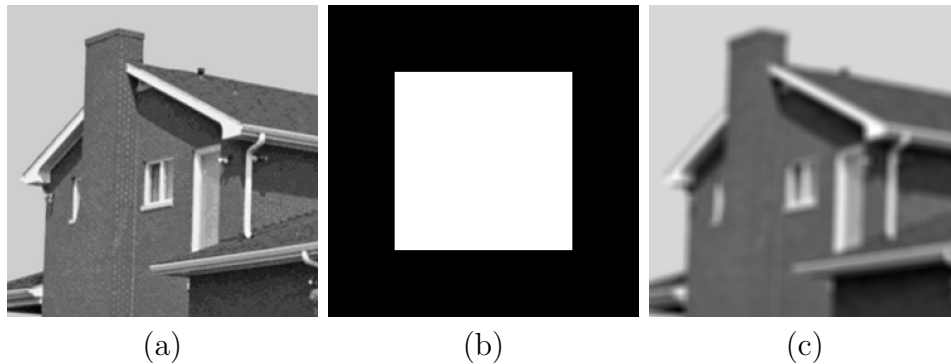
Figure 3: House test case: (a) true image (248×248 pixels), (b) PSF (14×14 pixels), (c) blurred and noise-free image.

5 synthetic test problems with different true images and apply the MAIT algorithm with the values of β of Table 5. For each test, we determine the minimum value of β for which (11) is satisfied for all k, average these values, and use the average value to solve the original problem.

Table 6 reports the RREs obtained with the AIT and MAIT methods for each noise level, as well as the average values of $\hat{\beta}$ computed by our scheme. We obtain similar results as in [8]. For $\sigma = 0.1\%$ the AIT method fails to converge, while the MAIT method converges and computes an accurate restoration. When the noise level is increased to 1% both the AIT and MAIT methods produce restorations of the same quality. Moreover, we can observe that our heuristic approach determines a suitable value of β. Figure 4 reports the evolution of the RRE versus the number of iterations. Figure 5 displays the computed solutions. Visual inspection of these figures are in agreement with Table 6.
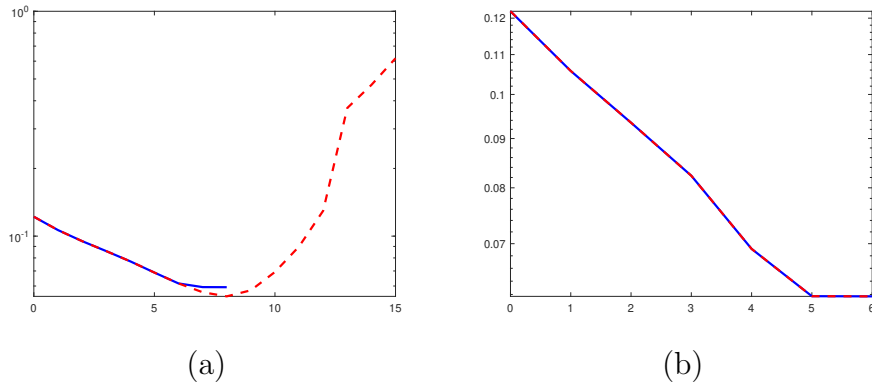
(a)   (b)

Figure 4: House test case evolution of the RRE throughout the iterations for the two considered noise levels. The red dashed line is refereed to AIT, while the solid blue line reports the RRE obtained with MAIT: (a) $\sigma = 0.1\%$, (b) $\sigma = 1\%$.
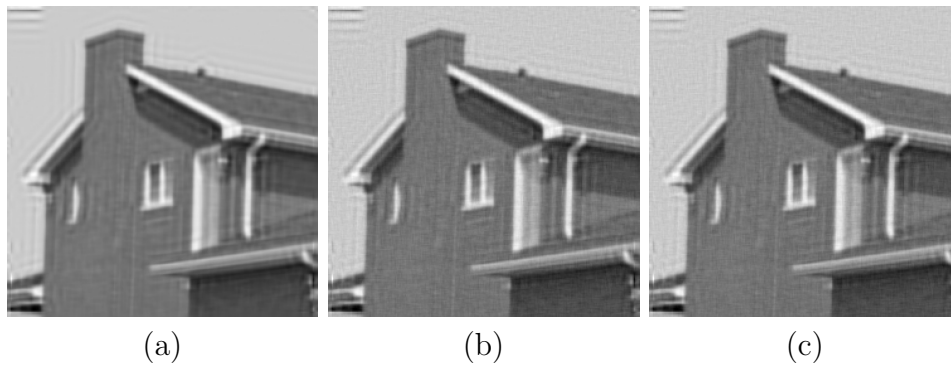

(a)   (b)   (c)

Figure 5: House test case reconstructions: (a) MAIT ($\sigma = 0.1\%$), (b) AIT ($\sigma = 1\%$), (c) MAIT ($\sigma = 1\%$).

5. Conclusions

In this work, we show that Assumption 2 and Assumption 1 are equivalents. This corrects a statement in [8]. Nevertheless, the theoretical results shown in [8] are correct. We propose a weaker assumption and show that under this new hypothesis the convergence properties of Algorithm 2 are maintained, provided that the initial approximate solution is not too far from the desired solution. Finally, we propose a heuristic approach for the automatic determination of the parameter β needed by Algorithm 2.

Acknowledgments

References

[1] P. C. Hansen, Rank Deficient and Discrete Ill-Posed Problems: Numerical Aspects of Linear Inversion, SIAM, Philadelphia, 1998.

[2] H. W. Engl, M. Hanke, A. Neubauer, Regularization of Inverse Problems, volume 375, Springer Science & Business Media, 1996.

[3] M. Hanke, C. W. Groetsch, Nonstationary iterated Tikhonov regularization, Journal of Optimization Theory and Applications 98 (1998) 37–53.

[4] S. C. Buranay, D. Subasi, O. C. Iyikal, On the two classes of high-order convergent methods of approximate inverse preconditioners for solving linear systems, Numerical Linear Algebra with Applications 24 (2017) e2111.

[5] S. C. Buranay, O. C. Iyikal, Approximate Schur-block ILU preconditioners for regularized solution of discrete ill-posed problems, Mathematical Problems in Engineering 2019 (2019) 1912535.

[6] S. C. Buranay, M. A. Özarslan, S. S. Falahhesar, Numerical solution of the Fredholm and Volterra integral equations by using modified Bernstein–Kantorovich operators, Mathematics 9 (2021) 1193.

[7] M. Donatelli, M. Hanke, Fast nonstationary preconditioned iterative methods for ill-posed problems, with application to image deblurring, Inverse Problems 29 (2013) 095008.

[8] A. Buccini, M. Donatelli, L. Reichel, W.-H. Zhang, A new nonstationary preconditioned iterative method for linear discrete ill-posed problems with application to image deblurring, Numerical Linear Algebra with Applications 28 (2021) e2353.

[9] A. Buccini, Regularizing preconditioners by non-stationary iterated Tikhonov with general penalty term, Applied Numerical Mathematics 116 (2017) 64–81.

[10] C. Garoni, S. Serra-Capizzano, Generalized locally Toeplitz sequences: Theory and Applications, volume 1, Springer, Cham, 2017.

[11] G. Barbarino, C. Garoni, S. Serra-Capizzano, Block generalized locally Toeplitz sequences: Theory and applications in the unidimensional case, Electronic Transactions on Numerical Analysis 53 (2020) 28–112.

[12] G. Barbarino, C. Garoni, S. Serra-Capizzano, Block generalized locally Toeplitz sequences: Theory and applications in the multidimensional case, Electronic Transactions on Numerical Analysis 53 (2020) 113–216.

[13] A. Buccini, L. Onisk, L. Reichel, A preconditioned Arnoldi–Tikhonov method for the solution of linear discrete ill-posed problems, Numerical Algorithms (in press).