A relative error tolerant iterated-Tikhonov method for solving ill-posed problems

A. Leitão[†‡§] A. L. Madureira[‡¶] B. F. Svaiter[‖]

February 28, 2024

Abstract

In this manuscript we propose and analyze an inexact iterated-Tikhonov method with relative error tolerance (ret-iT method) for obtaining, in a stable way, approximate solutions to linear ill-posed operator equations. Convergence analysis is provided. Numerical experiments are presented for an exponentially ill-posed elliptic problem, demonstrating significant improvement in performance compared to standard implementations of the iterated-Tikhonov (iT) method.

Keywords. Ill-posed problems; Iterated Tikhonov method; Inexact method; Relative error tolerance.

AMS Classification: 65J20, 47J06.

1 Introduction

The problem we are interested in consists of determining an unknown quantity $x^\star \in X$ from a set of data $y^\star \in Y$, where X, Y are Hilbert spaces. This is the tipical setting of *inverse problems* [1, 23, 26], where an unknown quantity of interest $x^\star$ must be determined, based on information obtained from some set of measured data.

In practical situations, we do not know the data $y^\star \in Y$ exactly. Instead, only approximate measured data $y^\delta \in Y$ satisfying

$$\|y^\delta - y^\star\| \ \le \ \delta \,, \tag{1.1}$$

is at hand. Here $\delta > 0$ represents the (known) level of noise, i.e. the accuracy of the measurements is known. The available noisy data $y^\delta \in Y$ are obtained by indirect measurements of the (unknown) parameter, this process is described by the mathematical model

$$Ax \ = \ y^\delta \,, \tag{1.2}$$

where $A : X \to Y$, is a bounded linear ill-posed operator, whose inverse $A^{-1} : Y \to X$ either do not exist, or is not continuous.

[†]EMAp, Getulio Vargas Fundation, Praia de Botafogo 190, 22250-900 Rio de Janeiro, Brazil
[‡]LNCC, Av. Getúlio Vargas 333, P.O. Box 95113, 25651-070 Petrópolis, Brazil
[§]On leave from Department of Mathematics, Federal Univ. of St. Catarina, P.O. Box 476, 88040-900 Floripa, Brazil
[¶]EPGE, Getulio Vargas Fundation, Praia de Botafogo 190, 22250-900 Rio de Janeiro, Brazil
[‖]IMPA, Estrada Dona Castorina 110, 22460-320 Rio de Janeiro, Brazil
Emails: acgleitao@gmail.com, alm@lncc.br, benar@impa.br.

The iterated-Tikhonov (iT) method

The starting point of our approach is the iterated-Tikhonov (iT) method [1, 8, 23]. The iT method is an efficient alternative for obtaining approximate solutions to the linear ill-posed problem (1.1), (1.2). The step of this iterative method reads

$$x^\delta_{\mathrm{iT,k}} \;:=\; \arg\min_x \left\{ \lambda_k \|Ax - y^\delta\|^2 + \|x - x^\delta_{\mathrm{iT,k-1}}\|^2 \right\}, \tag{1.3}$$

eq:it-exac

where $(\lambda_k) > 0$ is an appropriately chosen sequence of Lagrange multipliers [2]. This is equivalent to compute $x^\delta_{\mathrm{iT,k}} \in X$ such that

$$\lambda_k A^*(Ax^\delta_{\mathrm{iT,k}} - y^\delta) + x^\delta_{\mathrm{iT,k}} - x^\delta_{\mathrm{iT,k-1}} \;=\; 0\,, \tag{1.4}$$

eq:it-exac

here A^* is the adjoint operator to A. The literature on the iT method is extense and focus on distinct aspects, e.g., regularization properties [7, 12, 22, 21], rates of convergence [15, 28], *a posteriori* strategies for choosing the Lagrange multipliers [2], ciclic iT type methods [5].

An inexact iT method with relative error tolerance (ret-iT)

In this article we propose and analyze an inexact version of the iT method (1.3). The step of our iterative method is defined by relaxing (1.4) and computing an approximate solution to this equation (a relative error is allowed); see (2.3) for details.

The motivation for adopting this strategy is clear: computationaly, it is less expensive to obtain a solution to a relaxed problem than to calculate the exact solution (up to the computer precision) to the original problem.

We are able to estimate the progress towards the solution of the iterate in the ret-iT method, and compare this progress with the one obtained in the iT method (see Section 2). We observe that they exhibit comparable quality. The numerical findings presented in Section 5 support this conclusion.

Inexact Newton methods for the stable solution of nonlinear ill-posed problems have been considered in the literature (see, e.g., Rieder [25, 27] and Hanke [17, 18]). In all these approaches the criteria for computing an inexact Newton step are based on relative residual tolerance. The careful reader observes that the ret-iT method is not a reduction of those algorithms to the linear setting.

Outline of the manuscript

In Section 2 we define the inexact step for the iterated-Tikhonov type method considered in this article; see (2.3). Some preliminary inequalities are established and a gain estimate is derived (Lemma 2.4).

The relative error tolerant iterated-Tikhonov (ret-iT) method is presented in Section 3. First the exact data case is considered (see Algorithm 1); a monotonicity result is proved (Proposition 3.1) as well as a convergence result (Theorem 3.3). In the sequel, the noisy data case is addressed (Algorithm 2); finiteness of the stopping index is proved (Proposition 3.7) as well as monotonicity of the iteration error (Lemma 3.5). Under appropriate assumptions (Assumption 3.9, existence of an inner iteration) we prove stability and semiconvergence results for the proposed method (Theorems 3.10 and 3.11). In Section 4 we prove that the Conjugate Gradient (CG) method, combined with a particular stoping rule, satisfies Assumption 3.9. Section 5 is devoted to numerical experiments. The Inverse Potential

Problem (IPP) is used to test the efficiency of the proposed method. Conclusions and final remarks are presented in Section 6.

2 Defining the inexact step of the ret-iT method

sec:method-step

In this section we introduce the step of the *relative error tolerant inexact iterated-Tikhonov method* (ret-iT) considered in these notes.

We discuss first a single step of the proposed method. Let x_{k-1}^δ be the current iterate and $\lambda_k > 0$ be an appropriately choosen Lagrange multiplier. In the iT method (1.3), the next iterate is the solution of

$$\lambda_k A^*(Ax - y^\delta) + x - x_{k-1}^\delta \;=\; 0\,. \tag{2.1}$$

eq:it-g1

For our aims, it is convenient to rewrite this equation as

$$x \;=\; x_{k-1}^\delta - \lambda_k A^*(Ax - y^\delta)\,. \tag{2.2}$$

eq:it-g2

The method proposed in this work is based on a relaxation of equation (2.1) combined with a modified update rule. More precisely, given the current iterate $x_{k-1}^\delta \in X$, one computes an *auxiliary point* $\widetilde{x}_k^\delta \in X$ such that

def:step-inexact

$$\|\lambda_k A^*(A\widetilde{x}_k^\delta - y^\delta) + \widetilde{x}_k^\delta - x_{k-1}^\delta\| \;\le\; \sigma\|\widetilde{x}_k^\delta - x_{k-1}^\delta\|\,, \tag{2.3a}$$

eq:txk

where $0 \le \sigma < 1$. This auxiliary point $\widetilde{x}_k^\delta$ is used to define the next iterate $x_k^\delta \in X$ as

$$x_k^\delta \;:=\; x_{k-1}^\delta - \lambda_k A^*(A\widetilde{x}_k^\delta - y^\delta) \tag{2.3b}$$

eq:xk

(compare with (2.2)). Observe that, for $0 < \sigma < 1$, (2.3a) is a relaxation of (2.1) with *relative error tolerance*.

The motivation for the definition of the inexact step in (2.3) is twofold. First: it is very often easyer to obtain a solution $\widetilde{x}_k^\delta \in X$ of (2.3a), than to compute the exact solution of (2.1) (up to the computer precision) as in the iT method. Second: the progress towards the solution of the iterate in (2.3b) is quantitatively "almost as good" as the one obtained in the iT method (as shown in Lemma 2.4).

For the remaining of this section we consider the exact data case $y^\delta = y^\star$ (i.e. $\delta = 0$) and write x_k, $\widetilde{x}_k$ instead of x_k^δ, $\widetilde{x}_k^\delta$. For any $0 \le \sigma < 1$, the solution of (2.1) is also a solution (2.3a). Moreover, it is easy to verify that if either $Ax_{k-1} = y^\star$ or $\sigma = 0$, then the unique solution of (2.3a) is the unique solution of (2.1).

In what concerns the method we are proposing, the case of interest is $Ax_{k-1} \neq y^\star$ and $0 < \sigma < 1$. Any iterative method for solving (2.1) generates a sequence (z_j) which converges to the solution of this problem. In this case, whenever x_{k-1} is not already a solution of $Ax = y^\star$ and $0 < \sigma < 1$, the iterates $\widetilde{x}_k = z_j$ will eventually satisfy (2.3a), as shown in the next proposition.

Proposition 2.1. *Suppose that x_{k-1} is not a solution of $Ax = y^\star$, and let x^+ be the solution of (2.1), i.e.*

$$x^+ \;=\; (\lambda_k A^* A + I)^{-1}(x_{k-1} + \lambda_k A^* y^\star)\,.$$

For any $0 < \sigma < 1$ the solution set of (2.3a) contains x^+ in its interior.

Proof. Let

$$f(\widetilde{x}) := \|\lambda_k A^*(A\widetilde{x} - y^\star) + \widetilde{x} - x_{k-1}\| - \sigma\|\widetilde{x} - x_{k-1}\| \qquad (\widetilde{x} \in X).$$

The solution set of (2.3a) is the level set $f \leq 0$. It follows from the assumption $Ax_{k-1} \neq y^\star$ that $x^+ \neq x_{k-1}$; therefore, $f(x^+) = -\sigma\|x^+ - x_{k-1}\| < 0$. To end the proof, observe that f is continuous. □

Corollary 2.2. *Suppose that $Ax_{k-1} \neq y^\star$, $0 < \sigma < 1$. If (z_i) is a sequence in X which converges to the solution of (2.1), then for i large enough, $\widetilde{x}_k = z_i$ is a solution of (2.3a).*

In the next lemma, some basic inequalities relating x_{k-1}, $\widetilde{x}_k$ and x_k are established.

lem:prelim **Lemma 2.3.** *Let $\sigma \in [0, 1)$, $x_{k-1} \in X$. If $\widetilde{x}_k$, x_k satisfy (2.3a) and (2.3b), then:*

it:a *a)* $\|\widetilde{x}_k - x_k\| \leq \sigma\|\widetilde{x}_k - x_{k-1}\|$;

it:b *b)* $(1 - \sigma)\|\widetilde{x}_k - x_{k-1}\| \leq \|x_k - x_{k-1}\| \leq (1 + \sigma)\|\widetilde{x}_k - x_{k-1}\|$.

Proof. To prove item (a), substitute x_k with its definition at (2.3b) in $\|\widetilde{x} - x_k\|$ and use (2.3a). To prove item (b), observe that

$$\|\widetilde{x}_k - x_{k-1}\| - \|\widetilde{x}_k - x_k\| \leq \|x_k - x_{k-1}\| \leq \|\widetilde{x}_k - x_{k-1}\| + \|\widetilde{x}_k - x_k\|$$

and use item (a). □

Let $x^\star \in X$ be an exact solution of $Ax = y^\star$. In the next lemma we estimate the "gain" $\|x^\star - x_{k-1}\|^2 - \|x^\star - x_k\|^2$ obtained after a single step of the ret-iT method. The proof of this result will be simplifyed using the identity

$$\|a\|^2 - \|b\|^2 \;=\; \|a - c\|^2 - \|b - c\|^2 + 2\,\langle a - b, c\rangle\;, \tag{2.4}$$ eq:aux-bs

for a, b, $c \in X$. This identity will also be used in the next section.

m:gain-exact-new **Lemma 2.4.** *Let $k \geq 1$, $\lambda_k > 0$, $0 \leq \sigma < 1$ and $x_{k-1} \in X$. If $\widetilde{x}_k$ and x_k are as in (2.3a) and (2.3b), then*

$$\|x^\star - x_{k-1}\|^2 - \|x^\star - x_k\|^2 \;\geq\; 2\lambda_k\,\|A\widetilde{x}_k - y^\star\|^2 + (1 - \sigma^2)\,\|\widetilde{x}_k - x_{k-1}\|^2,$$

for any $x^\star$ solution of $Ax = y^\star$.

Proof. Using (2.4) for $a = x^\star - x_{k-1}$, $b = x^\star - x_k$ and $c = x^\star - \widetilde{x}_k$; together with (2.3b) we obtain

$$\begin{aligned}
\|x^\star - x_{k-1}\|^2 - \|x^\star - x_k\|^2 &= \|\widetilde{x}_k - x_{k-1}\|^2 - \|\widetilde{x}_k - x_k\|^2 \\
&\qquad + 2\,\langle x_k - x_{k-1}, x^\star - \widetilde{x}_k\rangle \\
&= \|\widetilde{x}_k - x_{k-1}\|^2 - \|\lambda_k A^*(A\widetilde{x}_k - y^\star) + \widetilde{x}_k - x_{k-1}\|^2 \\
&\qquad + 2\,\lambda_k\,\langle A^*(A\widetilde{x}_k - y^\star), \widetilde{x}_k - x^\star\rangle\,.
\end{aligned}$$

To end the proof, use (2.3a) to estimate the second norm at the right-hand side of the last equality and observe that $\langle A^*(A\widetilde{x}_k - y^\star), \widetilde{x}_k - x^\star\rangle = \|A\widetilde{x}_k - y^\star\|^2$. □

This lemma is a key result, from where many relevant consequences (e.g., monotonicity of the ret-iT method) can be derived, as we shall see in the next section.

3 The ret-iT method

In what follows we introduce and analyze the ret-iT method for computing stable approximate solutions to ill-posed problems of the form (1.1), (1.2). The exact data case ($\delta = 0$) is considered in Section 3.1, where our method is defined based on the inexact step in Section 2. In Section 3.2 the noisy data case is addressed.

3.1 The exact data case

The inexact step discussed in (2.3) leads us to the following conceptual algorithm:

[1] choose an initial guess $x_0 \in X$;

[2] choose $\sigma \in (0,1)$ and a sequence $(\lambda_k) > 0$;

[3] **for** $k \geq 1$ **do**

[3.1] find $\widetilde{x}_k \in X$ solution of (2.3a), i.e.

$$\|\lambda_k A^*(A\widetilde{x}_k - y^\star) + \widetilde{x}_k - x_{k-1}\| \ \leq \ \sigma \|\widetilde{x}_k - x_{k-1}\|;$$

[3.2] define the next iterate $x_k \in X$

$$x_k \ := \ x_{k-1} - \lambda_k A^*(A\widetilde{x}_k - y^\star);$$

end for

Algorithm 1: The ret-iT method in the exact data case.

Proposition 3.1. *Let $\sigma \in (0,1)$, $(\lambda_k) > 0$ and the sequences (x_k), $(\widetilde{x}_k)$ be defined as in Algorithm 1. The following assertions hold true:*

a) $\|\widetilde{x}_k - x_k\| \ \leq \ \sigma \|\widetilde{x}_k - x_{k-1}\|$.

b) $(1-\sigma)\|\widetilde{x}_k - x_{k-1}\| \leq \|x_k - x_{k-1}\| = \|\lambda_k A^*(A\widetilde{x}_k - y^\star)\| \leq (1+\sigma)\|\widetilde{x}_k - x_{k-1}\|$

c) *For any $x^\star \in X$ solution of $Ax = y^\star$,*

$$\|x^\star - x_{k-1}\|^2 - \|x^\star - x_k\|^2 \ \geq \ 2\lambda_k \, \|A\widetilde{x}_k - y^\star\|^2 + (1-\sigma^2)\, \|\widetilde{x}_k - x_{k-1}\|^2, \ \forall \ k \geq 1\,.$$

d) *The following series are summable:*

$$\sum_{k=0}^{\infty} \lambda_k \, \|A\widetilde{x}_k - y^\star\|^2, \quad \sum_{k=0}^{\infty} \|\widetilde{x}_k - x_{k-1}\|^2, \quad \sum_{k=0}^{\infty} \lambda_k^2 \, \|A^*(A\widetilde{x}_k - y^\star)\|^2. \tag{3.1}$$

e) *Adittionaly, if $(\lambda_k) \geq \lambda_{\min} > 0$, the following series are summable:*

$$\sum_{k=0}^{\infty} \|A\widetilde{x}_k - y^\star\|^2, \quad \sum_{k=0}^{\infty} \|A^*(A\widetilde{x}_k - y^\star)\|^2.$$

f) *The set $\mathcal{R}_k := \{z \in X;\ \|\lambda_k A^*(Az - y^\star) + z - x_{k-1}\| \leq \sigma \|z - x_{k-1}\|\}$ is uniformly bounded for all k ($\mathcal{R}_k$ is the solution set of the problem in Step [3.1]).*

Proof. Assertions (a), (b) and (c) were proved in Lemmata 2.3 and 2.4. In Assertion (d), the summability of the first two series in (3.1) follow from Assertion (c), a telescopic-sum argument, and $0 < \sigma < 1$. Moreover, Assertion (b) implies

$$(1+\sigma)\, \|\widetilde{x}_k - x_{k-1}\| \ \geq \ \|x_k - x_{k-1}\| \ = \ \lambda_k \, \|A^*(A\widetilde{x}_k - y^\star)\|\,,$$

from where the summability of the last series in (3.1) follows.

Assertion (e) follows from (d). Moreover, Assertion (f) follows from (c) (indeed, Assertion (c) implies $\|x^\star - x_0\|^2 \geq \|x^\star - x_{k-1}\|^2 \geq (1-\sigma^2)\|z - x_{k-1}\|^2$, for all $z \in S_k$). $\square$

rem:stationary **Remark 3.2.** *Notice that $Ax_k = y^\star$ for some $k \geq 0$ if and only if $A\widetilde{x}_{k+1} = y^\star$. In this case, both sequences become stationary, i.e. $x_j = \widetilde{x}_j = x_k$, for $j \geq k+1$.*

The proof the next theorem is based on the classical proof presented in [16, Theorem 2.3] using Cauchy sequence argument to establish convergence of the nonlinear Landweber iteration in the exact data case.

th:conv-exact **Theorem 3.3 (Convergence for exact data).** *Let (x_k) and $(\widetilde{x}_k)$ be sequences defined by Algorithm 1, with $\sigma \in (0,1)$ and $(\lambda_k) \geq \lambda_{\min} > 0$. Then (x_k) converges strongly to some $\bar{x} \in X$. Moreover, $\bar{x}$ is a solution of $Ax = y^\star$.*

The careful reader observes that the linear problem $Ax = y^\star$ admits a x_0-minimal norm solution, i.e. an element $x^\dagger \in X$ satisfying $Ax^\dagger = y^\star$ and $\|x^\dagger - x_0\| = \inf\{\|x - x_0\|;\ Ax = y^\star\}$ (see, e.g., [1, 8]). Moreover, $x^\dagger$ is the only solution of $Ax = y^\star$ with this property. On the other hand, from Step [3.2] of Algorithm 1 follows $x_{k+1} - x_k \in R(A^*) \subset N(A)^\perp$. An inductive argument shows that $\bar{x}$ in Theorem 3.3 satisfies $\bar{x} \in x_0 + N(A)^\perp$ and, consequently, $\bar{x} = x^\dagger$.

Proof. (of Theorem 3.3) We divide the proof in two separate cases:

Case I: $A\widetilde{x}_{k_0} = y^\star$ for some $k_0 \in \mathbb{N}$.
It follows from Remark 3.2, that $x_k = \widetilde{x}_k = \widetilde{x}_{k_0}$ for all $k > k_0$. Thus the strong convergence of (x_k) to $\bar{x} := \widetilde{x}_{k_0}$ (which, in this case, is a solution of $Ax = y^\star$) follows.

Case II: $A\widetilde{x}_k \neq y$ for all $k \in \mathbb{N}$.
In this case $(\|A\widetilde{x}_k - y^\star\|)$ is a strictly positive sequence. Moreover, it follows from Proposition 3.1 (e) that $\lim_k \|A\widetilde{x}_k - y^\star\| = 0$. Thus, there is a strictly monotone increasing sequence $(\ell_j) \in \mathbb{N}$ such that

$$\|A\widetilde{x}_k - y^\star\| \ \geq\ \|A\widetilde{x}_{\ell_j} - y^\star\|\,,\ 0 \leq k \leq \ell_j\,. \tag{3.2}$$

eq:sseq-mi

Notice that, given $k \geq 1$ and $z \in X$, it follows from (2.4) with $a = x_{k-1} - z$, $b = x_k - z$ and $c = \widetilde{x}_k - z$ that

$$\begin{aligned} \|x_{k-1} - z\|^2 - \|x_k - z\|^2 &= \|x_{k-1} - \widetilde{x}_k\|^2 - \|x_k - \widetilde{x}_k\|^2 + 2\,\langle x_{k-1} - x_k, \widetilde{x}_k - z\rangle \\ &\leq \|x_{k-1} - \widetilde{x}_k\|^2 + 2\,\langle x_{k-1} - x_k, \widetilde{x}_k - z\rangle\,. \end{aligned}$$

Thus, it follows from Step [3.2] of Algorithm 1

$$\begin{aligned} \|x_{k-1} - z\|^2 - \|x_k - z\|^2 &\leq \|x_{k-1} - \widetilde{x}_k\|^2 + 2\,\langle \lambda_k A^*(A\widetilde{x}_k - y^\star), \widetilde{x}_k - z\rangle\,. \\ &= \|x_{k-1} - \widetilde{x}_k\|^2 + 2\lambda_k\,\langle A\widetilde{x}_k - y^\star, A(\widetilde{x}_k - z)\rangle \\ &= \|x_{k-1} - \widetilde{x}_k\|^2 + 2\lambda_k\,\langle A\widetilde{x}_k - y^\star, (A\widetilde{x}_k - y^\star) + (y^\star - Az)\rangle \\ &\leq \|x_{k-1} - \widetilde{x}_k\|^2 + 2\lambda_k\Big[\|A\widetilde{x}_k - y^\star\|^2 + \|A\widetilde{x}_k - y^\star\|\|Az - y^\star\|\Big]. \end{aligned}$$

Thus, choosing $z = \widetilde{x}_{\ell_j}$ in the above inequality and arguing with (3.2) we obtain

$$\begin{aligned} &\|x_{k-1} - \widetilde{x}_{\ell_j}\|^2 - \|x_k - \widetilde{x}_{\ell_j}\|^2 \\ &\qquad \le \ \|\widetilde{x}_k - x_{k-1}\|^2 + 2\lambda_k \|A\widetilde{x}_k - y^\star\|^2 + 2\lambda_k \|A\widetilde{x}_k - y^\star\| \|A\widetilde{x}_{\ell_j} - y^\star\| \\ &\qquad \le \ \|\widetilde{x}_k - x_{k-1}\|^2 + 4\lambda_k \|A\widetilde{x}_k - y^\star\|^2 , \end{aligned}$$

for $k = 1, \dots, \ell_j$. Adding the above inequality for $k = m+1, m+2, \dots, \ell_j$ we conclude that

$$\|x_m - \widetilde{x}_{\ell_j}\|^2 \ \le \ \sum_{k=m+1}^{\ell_j} \Big[\|\widetilde{x}_k - x_{k-1}\|^2 + 4\lambda_k \|A\widetilde{x}_k - y^\star\|^2 \Big] \ \le \ s_m \, ,$$

where $s_m := \sum_{k=m}^{\infty} \big[\|\widetilde{x}_k - x_{k-1}\|^2 + 4\lambda_k \|A\widetilde{x}_k - y^\star\|^2 \big]$; notice that Proposition 3.1 (d) guarantees $\lim\limits_{m\to\infty} s_m = 0$.

Now, for $n > m$ we choose $\ell_j > n$ and estimate

$$\|x_n - x_m\| \ \le \ \|x_n - \widetilde{x}_{\ell_j}\| + \|\widetilde{x}_{\ell_j} - x_m\| \ \le \ \sqrt{s_n} + \sqrt{s_m} \ \le \ 2\sqrt{s_m}$$

(the sequence (s_m) is monotone decreasing). Thus, (x_k) is a Cauchy sequence and converges to some element $\bar{x} \in X$.

To prove that $\bar{x}$ is a solution of $Ax = y^\star$, it sufices to show that the residuals $\|Ax_k - y^\star\|$ converge to zero as $k \to \infty$. Since $\lim_k \|A\widetilde{x}_k - y^\star\| = \lim_k \|x_{k-1} - \widetilde{x}_k\| = 0$ (see Proposition 3.1 (d) and (e)), it follows that $\lim_k \|Ax_k - y^\star\| = 0$ concluding the proof. □

3.2 The noisy data case

sec:method-noise

In the noisy data case, $\delta > 0$, the step of the ret-iT method is defined by (2.3). Based by the inexact step in (2.3a), (2.3b) we propose in Algorithm 2 the ret-iT method for the noisy data case.

[1] choose an initial guess $x_0 \in X$; set $k := 1$; set $x_0^\delta = x_0$;

[2] choose constants $\sigma \in (0, 1)$, $\tau > 1$ and a sequence $(\lambda_k) > 0$;

[3] **repeat**

  [3.1] compute $\widetilde{x}_k^\delta \in X$ as a solution of (2.3a), i.e.

$$\|\lambda_k A^*(A\widetilde{x}_k^\delta - y^\delta) \ + \ \widetilde{x}_k^\delta - x_{k-1}^\delta\| \ \le \ \sigma \|\widetilde{x}_k^\delta - x_{k-1}^\delta\|;$$

  [3.2] **if** $\big(\|A\widetilde{x}_k^\delta - y^\delta\| > \tau\delta\big)$ **then**

    $x_k^\delta \ := \ x_{k-1}^\delta - \lambda_k A^*(A\widetilde{x}_k^\delta - y^\delta)$;

    $k \ := \ k + 1$;

   else

    EXIT LOOP;

   end if

  end repeat

[4] $k^*(\delta) \ := \ k - 1$;

alg:IIT-method

Algorithm 2: The ret-iT method in the noisy data case.

Observe that Algorithm 2 generates sequences $(x_k^\delta)_{k=0}^{k^*}$ and $(\widetilde{x}_k^\delta)_{k=1}^{k^*+1}$. Some relevant remarks folllow:

— The iterates $\widetilde{x}_1^\delta, \ldots, \widetilde{x}_{k^*+1}^\delta$ (and $x_1^\delta, \ldots, x_{k^*}^\delta$) are computed solving in each iteration a feasible problem.

— If $\delta = 0$, Algorithm 2 reduces to the ret-iT method for exact data.

— The stopping criterion used in Algorithm 2 (see Step [3.2]) is based on the discrepancy principle applied to $\widetilde{x}_k^\delta$, i.e. the iteration is stopped at step $k^* = k^*(\delta) \in \mathbb{N}$ satisfying

$$k^* \ := \ \max\{k \in \mathbb{N};\ \|A\widetilde{x}_j^\delta - y^\delta\| > \tau\delta,\ j = 1, \ldots, k\}.$$

— For $k^* \in \mathbb{N}$ defined in Step [4] it holds $\|A\widetilde{x}_{k^*}^\delta - y^\delta\| > \tau\delta$ as well as $\|A\widetilde{x}_{k^*+1}^\delta - y^\delta\| \le \tau\delta$.

Remark 3.4. *A result analog to the one stated in Proposition 2.1 holds true in the noisy data case. Notice that if $\widetilde{x} \in X$ satisfies $\|A\widetilde{x} - y^\delta\| > \delta$, then $A\widetilde{x} - y^\delta \notin Ker(A^*)$. Thus, arguing as in the proof of Proposition 2.1 we conclude that if $\|Ax_{k-1}^\delta - y^\delta\| > \delta$, x_+^δ is the solution of (2.1) and $\sigma \in (0,1)$ then the solution set of (2.3a) contains x_+^δ in its interior.*

Consequently, a conclusion analog to the one discussed in Corollary 2.2 holds true in the noisy data case.

Monotonicity results

In the sequel we establish a result, which is analog to the one discussed in Lemma 2.4.

lem:gain-noise

Lemma 3.5. *Let $\lambda_k > 0$ and $0 \le \sigma < 1$. Given $x_{k-1}^\delta \in X$, let $\widetilde{x}_k^\delta$ and x_k^δ be defined as in (2.3a), (2.3b) respectively. If $\|A\widetilde{x}_k^\delta - y^\delta\| > \delta$, then for any $x^\star \in X$ solution of $Ax = y$ it holds*

$$\|x^\star - x_{k-1}^\delta\|^2 - \|x^\star - x_k^\delta\|^2 \ge \lambda_k \left[\|A\widetilde{x}_k^\delta - y^\delta\|^2 - \delta^2 + \|A\widetilde{x}_k^\delta - y\|^2\right] + (1 - \sigma^2)\, \|\widetilde{x}_k^\delta - x_{k-1}^\delta\|^2.$$

Consequently, $\|x^\star - x_{k-1}^\delta\|^2 - \|x^\star - x_k^\delta\|^2 \ge \lambda_k \left[\|A\widetilde{x}_k^\delta - y^\delta\|^2 - \delta^2\right]$.

Proof. Due to (2.3) it holds $\|\widetilde{x}_k^\delta - x_k^\delta\| \le \sigma \|\widetilde{x}_k^\delta - x_{k-1}\|^\delta$. From this inequality together with (2.4) for $a = x^\star - x_{k-1}^\delta$, $b = x^\star - x_k^\delta$ and $c = x^\star - \widetilde{x}_k^\delta$ we obtain

$$\|x^\star - x_{k-1}^\delta\|^2 - \|x^\star - x_k^\delta\|^2 \ge 2\,\langle x_k^\delta - x_{k-1}^\delta, x^\star - \widetilde{x}_k^\delta\rangle + (1 - \sigma^2)\, \|\widetilde{x}_k^\delta - x_{k-1}^\delta\|^2.$$

Thus, to prove the lemma it suffices to prove that

$$\langle x_k^\delta - x_{k-1}^\delta, x^\star - \widetilde{x}_k^\delta\rangle \ \ge \ \tfrac{1}{2}\lambda_k \left[\|A\widetilde{x}_k^\delta - y^\delta\|^2 - \delta^2 + \|A\widetilde{x}_k^\delta - y\|^2\right].$$

Define the quadratic functional $f_\delta(x) := \frac{1}{2}\|Ax - y^\delta\|^2$. Notice that $f_\delta(\widetilde{x}_k^\delta) = \frac{1}{2}\|A\widetilde{x}_k^\delta - y^\delta\|^2$, $\nabla f_\delta(\widetilde{x}_k^\delta) = \lambda_k^{-1}(x_{k-1}^\delta - x_k^\delta)$ and $f_\delta(x^\star) \le \frac{1}{2}\delta^2$. Therefore, it follows from

$$f_\delta(x^\star) \ = \ f_\delta(\widetilde{x}_k^\delta) \ + \ \langle \nabla f_\delta(\widetilde{x}_k^\delta), x^\star - \widetilde{x}_k^\delta\rangle \ + \ \tfrac{1}{2}\langle (x^\star - \widetilde{x}_k^\delta), A^*A(x^\star - \widetilde{x}_k^\delta)\rangle$$

that $\frac{1}{2}\delta^2 \ge \frac{1}{2}\|A\widetilde{x}_k^\delta - y^\delta\|^2 + \langle \lambda_k^{-1}(x_{k-1}^\delta - x_k^\delta), x^\star - \widetilde{x}_k^\delta\rangle + \frac{1}{2}\|A(x^\star - \widetilde{x}_k^\delta)\|^2$. Therefore,

$$\lambda_k^{-1}\,\langle x_k^\delta - x_{k-1}^\delta, x^\star - \widetilde{x}_k^\delta\rangle \ \ \ge \ \ \tfrac{1}{2}\|A\widetilde{x}_k^\delta - y^\delta\|^2 - \tfrac{1}{2}\delta^2 + \tfrac{1}{2}\|A\widetilde{x}_k^\delta - y\|^2,$$

concluding the proof. $\square$

The next result is an imediate consequence of Lemma 3.5 and (2.3a); it will be used in the proof of the stability Theorem 3.10.

cor:new **Corollary 3.6.** *Under the assumptions of Lemma 3.5, for all* $0 \le k \le k^*(\delta)$,

$$\|x^\star - x_k^\delta\| \ \le \ \|x^\star - x_0\|$$

and for all $1 \le k \le k^*(\delta) + 1$, $\|x^\star - \widetilde{x}_k^\delta\| \le \big(\lambda_k \|A\|^2 \|x_0 - x^\star\| + \delta\big)(1-\sigma)^{-1}$.

The stopping index

In what follows we establish the finiteness of the stopping index k^* defined in Section 3.2. The next result is a direct consequence of Lemma 3.5.

prop:k-finite **Proposition 3.7.** *Let the sequences* (x_k^δ) *and* $(\widetilde{x}_k^\delta)$, *and* $k^* \in \mathbb{N}$ *be defined by Algorithm 2 with* $\tau > 1$, $\sigma \in (0,1)$ *and* $(\lambda_k) > 0$. *If* $\sum_k \lambda_k = \infty$ *then Algorithm 2 stops after a finite number of steps* $k^*(\delta) \in \mathbb{N}$.

Additionaly, if $(\lambda_k) \ge \lambda_{\min} > 0$, *then*

$$k^*(\delta) \ \le \ \|x^\star - x_0^\delta\|^2 \left(\lambda_{\min}(\tau^2 - 1)\delta^2\right)^{-1}.$$

Proof. Adding up the inequality in Lemma 3.5 for $k = 1, \ldots, k^*$, and observing Step [3.2] of Algorithm 2, we derive the estimate

$$\|x^\star - x_0^\delta\|^2 \ \ge \ \sum_{k=1}^{k^*} \lambda_k \big[\|A\widetilde{x}_k^\delta - y^\delta\|^2 - \delta^2\big] \ \ge \ (\tau^2 - 1)\delta^2 \sum_{k=1}^{k^*} \lambda_k \,,$$

from where the finiteness of k^* follows.

Additionaly, if the assumption $(\lambda_k) \ge \lambda_{\min} > 0$ holds, we derive from the last inequality the estimate $\|x^\star - x_0^\delta\|^2 \ge \lambda_{\min}(\tau^2 - 1)\delta^2\, k^*$, concluding the proof. $\square$

rem:stop-k **Remark 3.8.** *Step [3.2] of Algorithm 2 allow us to detect the first indek* $k \ge 1$ *such that the corresponding* $\widetilde{x}_k^\delta \in X$ *satisfies* $\|A\widetilde{x}_k^\delta - y^\delta\| \le \tau\delta$. *Notice that this* $\widetilde{x}_k^\delta$ *is not used to compute* x_{k+1}^δ. *Instead, the iteration is terminated at* $x_{k^*}^\delta$ *with* $k^* := k - 1$.

The reason for this choice becomes evident from Lemma 3.5. Namely, the "gain inequality" *in this corollary holds only if* $\widetilde{x}_k^\delta \in X$ *obtained in Step [3.2] satisfies* $\|A\widetilde{x}_k^\delta - y^\delta\| > \tau\delta$.

The careful reader notices that it would be possible to terminate the ret-iT method with $k^* := k$ *in Step [4] and* $x_k^\delta := \widetilde{x}_k^\delta$, *since* $\widetilde{x}_k^\delta$ *is the first point produced by Algorithm 2 which belongs to the level-set* $\{x \in X;\ \|Ax - y^\delta\| \le \tau\delta\}$.

Stability and semiconvergence

In the noisy data case, to ensure stability and semiconvergence we need additional assumptions on how $\widetilde{x}_k^\delta$ is computed; next we discuss and formalize these assumptions.

We will suppose that an "inner" iterative procedure is used to compute $\widetilde{x}_k^\delta$ for all k, choosing it as the first iterate which satisfies the error criterion (2.3a). That is, given λ_k, x_{k-1}^δ and y^δ, an iterative procedure generates a sequence

$$(z_n)_{n=0,1,\ldots} \ = \ (z_{n,\lambda_k,x_{k-1}^\delta,y^\delta})_{n=0,1,\ldots} \quad \text{with} \quad z_{0,\lambda_k,x_{k-1}^\delta,y^\delta} = x_{k-1}^\delta \,,$$

which converges to the exact solution of $\lambda_k A^*(Az - y^\delta) + z - x^\delta_{k-1} = 0$; the next "outer" iterate $\widetilde{x}^\delta_k$ is the first element genarated by this iterative procedure that satisfies the error criterion

$$\widetilde{x}^\delta_k = z_{n^\delta_k} \text{ where } n^\delta_k = \min\{n : \|\lambda_k A^*(Az_n - y^\delta) + z_n - x^\delta_{k-1}\| \leq \sigma \|z_n - x^\delta_{k-1}\|\}.$$

We make two aditional assumptions:
— for each n, the iterate z_n depends continuously on x^δ_{n-1} and y^δ;
— for each k, the number of inner steps n^δ_k is uniformly bounded for $0 < \delta < \overline{\delta}$ (where $\overline{\delta} > 0$ is fixed).

These assumptions are formalized below.

ass:method **Assumption 3.9.** *In Step [3.1] of Algorithm 2, each $\widetilde{x}^\delta_k$ is computed by an* **"inner" iterative method** *whose iterates are modeled as a family of mappings on $(x^\delta_{k-1},\, y^\delta)$*

$$T_{n,\lambda_k} : X \times Y \to X \qquad (n = 0, 1, \dots) \tag{3.3}$$ def:Tn

with the following properties:
1) $T_{0,\lambda_k}(x^\delta_{k-1}, y^\delta) = x^\delta_{k-1}$;
2) For each n and λ_k, the mapping T_{n,λ_k} is continuous;
3) for each $1 \leq k \leq k^*(\delta) + 1$, there exists $n^\delta_k \in \mathbb{N}$ such that

$$\widetilde{x}^\delta_k = T_{n^\delta_k,\lambda_k}(x^\delta_{k-1}, y^\delta).$$

Additionaly, let $(\delta_j) \in \mathbb{R}$ be a given zero sequence, and $(y^{\delta_j}) \in Y$ a corresponding sequence of noisy data satisfying (1.1); and let $(x^{\delta_j}_k)$ and $(\widetilde{x}^{\delta_j}_k)$ be (finite) sequences generated by Algorithm 2, for each $j \in \mathbb{N}$.
4) For each $k \in \mathbb{N}$, $\widetilde{x}^{\delta_j}_k$ is generated by the "inner" iterative method with **at most** *N_k steps, i.e. N_k does not depend on j. That is, for each $k \in \mathbb{N}$*

$$N_k \;=\; sup\{n^{\delta_j}_k : for\ j = 1, 2, \dots\ with\ k^*(\delta_j) \geq k\} \;<\; \infty.$$

We are now ready to state and prove a stability result for the ret-iT method.

th:stabil **Theorem 3.10** (Stability)**.** *Let (δ_j) be a zero sequence and $(y^{\delta_j}) \in Y$ a corresponding sequence of noisy data satisfying (1.1). For each $j \in \mathbb{N}$, let $\big(x^{\delta_j}_k\big)_{k=0}^{k^*(\delta_j)}$ and $\big(\widetilde{x}^{\delta_j}_k\big)_{k=1}^{k^*(\delta_j)+1}$ be finite sequences generated by Algorithm 2.*
If Assumption 3.9 holds, then there exist $K^* \in \mathbb{N} \cup \{\infty\}$, a subsequence of (δ_j) (denoted again by (δ_j)), and a pair of sequences (x_k) and $(\widetilde{x}_k)$ generated by Algorithm 1; such that $x^{\delta_j}_k \to x_k$, $\widetilde{x}^{\delta_j}_{k+1} \to \widetilde{x}_{k+1}$, as $j \to \infty$, for all $k \in \mathbb{N}$ with $k \leq K^*$.

Proof. There exists a subsequence (denoted again by (δ_j)) s.t.

$$k^*(\delta_1) \leq k^*(\delta_2) \leq k^*(\delta_3) \leq \dots$$

Denote

$$K^* = \lim_{j\to\infty} k^*(\delta_j), \qquad K^* \in \mathbb{N} \cup \{\infty\},$$

and let $J_k \subset \mathbb{N}$ be the set of indices $j \in \mathbb{N}$ for which $x_k^{\delta_j}$ is defined, i.e.

$$J_k = \{j \in \mathbb{N} \,:\, k^*(\delta_j) \geq k\} \qquad (1 \leq k \leq K^* + 1).$$

Note that each J_k is an unbounded set of consecutive natural numbers.

Fix $1 \leq k \leq K^* + 1$. It follows from Assumption 3.9, item 3, and Corollary 3.6 that $(n_k^{\delta_j})_{j \in J_k}$, as defined in Assumption 3.9, item 2, is bounded. Using a diagonal process for subsequence extraction, we conclude that there exists a subsequence again denoted by (δ_j) and a sequence (finite or otherwise) $(n_k)_{\{k \in \mathbb{N}\,:\,1 \leq k \leq K^*+1\}}$ such that

$$n_k^{\delta_j} = n_k, \qquad (1 \leq k \leq K^* + 1 \text{ and } j \in J_k). \tag{3.4}$$

eq:aux1

Observe that

$$\widetilde{x}_k^{\delta_j} = T_{n_k, \lambda_k}(x_{k-1}^{\delta_j}, y^{\delta_j}) \qquad (1 \leq k \leq K^* + 1 \text{ and } j \in J_k). \tag{3.5}$$

eq:aux2

We claim that for any $1 \leq k \leq K^* + 1$,

$$\exists \lim_{j \to \infty} x_{k-1}^{\delta_j}, \qquad \exists \lim_{j \to \infty} \widetilde{x}_k^{\delta_j}.$$

For $k = 1$, in view of Step [1], the first above limit exists. Suppose that the first above limit exists for some $1 \leq k \leq K^* + 1$. Since T_{n_k, λ_k} is continuous, it follows from (3.5) (and the assumption $\delta_j \to 0$) that the second limit also exists. If $k < K^*$, then the first above limit also exists for $k' = k + 1 \leq K^*$, because $k' - 1 = k$ and $x_{k'}^{\delta_j}$ depends continuously on $x_{k'-1}^{\delta_j}$, $\widetilde{x}_{k'}^{\delta_j}$ and y^{δ_j}.

Let

$$x_k = \lim_{j \to \infty} x_k^{\delta_j} \quad 1 \leq k \leq K^*, \qquad \widetilde{x}_k = \lim_{j \to \infty} \widetilde{x}_k^{\delta_j} \quad 1 \leq k \leq K^* + 1. \tag{3.6}$$

In view of Algorithm 2 we have

$$\begin{aligned} &\|\lambda_k A^*(A\widetilde{x}_k^{\delta_j} - y^{\delta_j}) + \widetilde{x}_k^{\delta_j} - x_{k-1}^{\delta_j}\| \leq \sigma \|\widetilde{x}_k^{\delta_j} - x_{k-1}^{\delta_j}\| && 1 \leq k \leq K^* + 1; \\ &x_k^{\delta_j} = x_{k-1}^{\delta_j} - \lambda_k A^*(A\widetilde{x}_k^{\delta_j} - y^{\delta_j}) && 1 \leq k \leq K^*. \end{aligned}$$

Therefore, taking the limit $j \to \infty$ we conclude that

$$\|\lambda_k A^*(A\widetilde{x}_k - y) + \widetilde{x}_k - x_{k-1}\| \leq \sigma \|\widetilde{x}_k - x_{k-1}\| \qquad 1 \leq k \leq K^* + 1; \tag{3.7a}$$

$$x_k = x_{k-1} - \lambda_k A^*(A\widetilde{x}_k - y) \qquad 1 \leq k \leq K^*. \tag{3.7b}$$

$\square$

We conclude this section addressing a regularization property of the ret-iT method.

th:semiconv

Theorem 3.11 (Semi-convergence). *Let (δ_j) be a zero sequence and $(y^{\delta_j}) \in Y$ a corresponding sequence of noisy data satisfying (1.1). For each $j \in \mathbb{N}$, let $(x_k^{\delta_j})$ and $(\widetilde{x}_k^{\delta_j})$, for $0 \leq k \leq k^*(\delta_j)$, be sequences defined by Algorithm 2.*

If Assumption 3.9 holds, then $(x_{k^*(\delta_j)}^{\delta_j})_j$ and $(\widetilde{x}_{k^*(\delta_j)+1}^{\delta_j})_j$ converge strongly to $x^\dagger$, the x_0-minimal norm *solution of $Ax = y$.*

Proof. It suffices to prove that every subsequence of $(x^{\delta_j}_{k^*(\delta_j)})_j$ has itself a subsequence converging strongly to $x^\dagger$, the same holding for $(\widetilde{x}^{\delta_j}_{k^*(\delta_j)+1})_j$.

We denote an arbitrary subsequence of (δ_j) again by (δ_j). Two distinct cases must be considered, depending on the corresponding subsequence $\big(k^*(\delta_j)\big)_j \in \mathbb{N}$.

Case I: The subsequence $\big(k^*(\delta_j)\big)_j$ is bounded.
Notice that $\big(k^*(\delta_j)\big)_j$ is a bounded sequence of natural numbers. Therefore there exists $K \in \mathbb{N}$ and a subsequence (δ_{j_m}) of (δ_j) such that

$$k^*(\delta_{j_m}) = K \qquad m = 1, 2, \ldots$$

It follows from Theorem 3.10 –applied to (δ_{j_m}), $(y^{\delta_{j_m}})$, $(x_k^{\delta_{j_m}})_{k=0}^{k^*(\delta_{j_m})}$, $(\widetilde{x}_k^{\delta_{j_m}})_{k=1}^{k^*(\delta_{j_m})+1}$– that there exists a subsequence of (δ_{j_m}) –denoted again by (δ_{j_m})– and a pair of sequences (x_k) and $(\widetilde{x}_k)$ generated by Algorithm 1, such that $x_k^{\delta_{j_m}} \to x_k$, $\widetilde{x}_{k+1}^{\delta_{j_m}} \to \widetilde{x}_{k+1}$, as $j_m \to \infty$, for $k = 1, \ldots\ K$.

We claim that $\widetilde{x}_{K+1}$ is a solution of $Ax = y$. Indeed, since $k^*(\delta_{j_m}) = K$ for all indices j_m, we have $\|A\widetilde{x}_{K+1}^{\delta_{j_m}} - y^{\delta_{j_m}}\| \le \tau\delta_{j_m}$ (see Step [3.2] of Algorithm 2). Thus,

$$\|A\widetilde{x}_{K+1} - y\| \;\le\; \lim_{j_m\to\infty} \big[\|A\widetilde{x}_{K+1}^{\delta_{j_m}} - y^{\delta_{j_m}}\| + \|y^{\delta_{j_m}} - y\|\big] \;\le\; \lim_{j_m\to\infty} (\tau+1)\delta_{j_m} \;=\; 0$$

and $A\widetilde{x}_{K+1} = y$. Now, it follows from Remark 3.2 that $x_K = \widetilde{x}_{K+1}$. By Algorithm 2, $x_K - x_0$ is in the range of A^*, concluding the proof of Case I.

Case II: The subsequence $\big(k^*(\delta_j)\big)_j$ is not bounded.
In this case, there exists a *monotone increasing* (sub)subsequence $k^*(\delta_1) \le k^*(\delta_2) \le \ldots$, again denoted by $\big(k^*(\delta_j)\big)_j$, such that $k^*(\delta_j) \to \infty$ as $j \to \infty$.

Notice that the subsequence (δ_j) and the corresponding sequences (y^{δ_j}), $(x_k^{\delta_j})$ and $(\widetilde{x}_k^{\delta_j})$ satisfy the assumptions of Theorem 3.10. Denote by (δ_j), (x_k) and $(\widetilde{x}_k)$ the subsequence and sequences specified in the conclusion part of that theorem. In particular $x_k^{\delta_j} \to x_k$ as $j \to \infty$.

Fix $\varepsilon > 0$. From Theorem 3.3 we know that $x_k \to x^\dagger$ as $k \to \infty$; hence, there exists $K_\varepsilon \in \mathbb{N}$ s.t. $\|x_k - x^\dagger\| \le \frac{1}{2}\varepsilon$ for $k \ge K_\varepsilon$. On the other hand, from the choice of $\big(k^*(\delta_j)\big)_j$, follows the existence of $J \in \mathbb{N}$ s.t. $k^*(\delta_j) \ge K_\varepsilon$, for all $j \ge J$. Thus, From Lemma 3.5 follows

$$\|x^{\delta_j}_{k^*(\delta_j)} - x^\dagger\| \;\le\; \|x^{\delta_j}_{K_\varepsilon} - x^\dagger\|\,,\ \forall\, j \ge J\,.$$

As $x_k^{\delta_j} \to x_k$ as $j \to \infty$, there exists $L \in \mathbb{N}$, $L > J$, such that

$$\|x^{\delta_j}_{K_\varepsilon} - x_{K_\varepsilon}\| \;\le\; \tfrac{1}{2}\varepsilon\,,\ \forall\, j \ge L\,.$$

Thus, for $j \ge L$ follows

$$\|x^{\delta_j}_{k^*(\delta_j)} - x^\dagger\| \;\le\; \|x^{\delta_j}_{K_\varepsilon} - x^\dagger\| \;\le\; \|x^{\delta_j}_{K_\varepsilon} - x_{K_\varepsilon}\| + \|x_{K_\varepsilon} - x^\dagger\| \;\le\; \varepsilon\,.$$

Since $\varepsilon > 0$ is arbitrary, it follows that $x^{\delta_j}_{k^*(\delta_j)} \to x^\dagger$, as $j \to \infty$ concluding the proof of Case II. ∎

4 The CG method as inner iteration

sec:cg

To prove stability and semiconvergence of Algorithm 2 (Theorems 3.10 and 3.11) we required Assumption 3.9. This assumption concerns existence and properties of an inner iterative method for the computation of $\tilde{x}_k^\delta$, $k = 1, \dots, k^*(\delta) + 1$, as specified in Step [3.1] of Algorithm 2. We will prove that the Conjugate Gradient (CG) method [11, 13, 8, 14] combined with a particular stopping rule, as this inner iterative method, satisfies Assumption 3.9.

We begin by presenting in Algorithm 3 the CG method with a relative error stopping rule for finding $\Delta x_k \in X$ such that $\tilde{x}_k^\delta = x_{k-1}^\delta + \Delta x_k$ satisfies (2.3a). Fix $k \geq 1$ and let

$$Q_k := \lambda_k A^* A + I \quad \text{and} \quad b_k := \lambda_k A^* y^\delta + x_{k-1}^\delta \,. \tag{4.1}$$

def:qb

Notice that $\tilde{x}_k^\delta$ is an approximate solution of the linear equation $Q_k x - b_k = 0$ satisfying the relative error tolerance

$$\| Q_k \, \tilde{x}_k^\delta - b_k \| \;\leq\; \sigma \| \tilde{x}_k^\delta - x_{k-1}^\delta \| \tag{4.2}$$

eq:probl_s

(compare with (2.3a)). Observe that the norm of that linear equation's residual at $\tilde{x}_k^\delta$ is at the left hand side of (4.2). Thus, we use this inequality as stopping rule of the proposed variant of the CG method. Recall that $\sigma \in (0, 1)$.

Remark 4.1. *Equivalently, $\Delta x_k = \tilde{x}_k^\delta - x_{k-1}^\delta$ satisfies*

$$\| Q_k \, \Delta x_k - c_k \| \;\leq\; \sigma \| \Delta x_k \| \,, \tag{4.3}$$

eq:probl_s

with $c_k = \lambda_k A^*(y^\delta - A x_{k-1}^\delta)$. I.e. Δx_k is an approximate solution of $Q_k x = c_k$ satisfying the relative error tolerance (4.3).

[1] $n := 0$; $s_0 := 0$; $r_0 := c_k$; $p_0 = r_0$;

[2] **if** $\big(r_0 = 0\big)$ **then** EXIT **end if**;

[3] **repeat**

  $\alpha_n := \langle r_n, r_n \rangle / \langle p_n, Q_k p_n \rangle$;

  $s_{n+1} := s_n + \alpha_n \, p_n$;

  $r_{n+1} := r_n - \alpha_n Q_k p_n$;

  if $\big(\|r_{n+1}\| > \sigma \|s_{n+1}\|\big)$ **then**

    $p_{n+1} := r_{n+1} + \frac{\langle r_{n+1}, r_{n+1} \rangle}{\langle r_n, r_n \rangle} \, p_n$;

    $n := n + 1$;

  else

    GO TO [4];

  end if

end repeat

[4] $n^\ddagger := n + 1$; $\Delta x_k = s_{n^\ddagger}$; $\tilde{x}_k^\delta := x_{k-1}^\delta + \Delta x_k$; EXIT

alg:CG-method

Algorithm 3: CG method for solving (4.3) with relative error stopping rule.

Direct use of (4.1) shows that for all $x \in X$,

$$\|x\|^2 \;\le\; \langle x, Q_k x\rangle \;\le\; (\lambda_k \|A\|^2 + 1)\|x\|^2 .$$

Therefore, using the notation $\operatorname{cond}(Q_k)$ for the condition number of Q_k, we have

$$\operatorname{cond}(Q_k) \;\le\; \lambda_k \|A\|^2 + 1\,.$$

The norm on X induced by Q_k is defined as

$$\|x\|_{Q_k} := \sqrt{\langle x, Q_k x\rangle},\;\; x \in X\,. \tag{4.4}$$

def:normQ

First we prove finite termination of Algorithm 3.

lem:fin-term

Lemma 4.2 (Finite termination)**.** *Algorithm 3 has finite termination and*

$$n^{\ddagger} \;\le\; \left\lceil \ln\Big(\frac{(\lambda_k\|A\|^2+1)+\sigma\sqrt{\lambda_k\|A\|^2+1}}{\sigma/2}\Big)\,\Big[\ln\Big(\frac{\sqrt{\kappa}+1}{\sqrt{\kappa}-1}\Big)\Big]^{-1}\right\rceil \tag{4.5}$$

eq:nddag-e

where $\kappa := \operatorname{cond}(Q_k)$.[1]

Proof. Notice that $A \neq 0$ is ill-posed and $1 < \kappa \le \lambda_k\|A\|^2 + 1$. Consequently, $(\sqrt{\kappa}+1)(\sqrt{\kappa}-1)^{-1} > 1$, and there exists a smallest $n_0 \in \mathbb{N}$ such that

$$\frac{1}{2}\Big(\frac{\sqrt{\kappa}+1}{\sqrt{\kappa}-1}\Big)^{n_0} \;\ge\; \frac{(\lambda_k\|A\|^2+1)+\sigma\sqrt{\lambda_k\|A\|^2+1}}{\sigma} \;>\; 1\,, \tag{4.6}$$

eq:nddag-e

i.e.

$$n_0 \;=\; \left\lceil \ln\Big(\frac{(\lambda_k\|A\|^2+1)+\sigma\sqrt{\lambda_k\|A\|^2+1}}{\sigma/2}\Big)\,\Big[\ln\Big(\frac{\sqrt{\kappa}+1}{\sqrt{\kappa}-1}\Big)\Big]^{-1}\right\rceil . \tag{4.7}$$

eq:nddag-e

Since $s_0 = 0$ in Algorithm 3, it follows from Lemma A.3 for $n = n_0$, $Q = Q_k$ and $c = c_k$ that

$$\frac{\|s^+\|_{Q_k}}{\|s_{n_0} - s^+\|_{Q_k}} \;\ge\; \frac{1}{2}\Big(\frac{\sqrt{\kappa}+1}{\sqrt{\kappa}-1}\Big)^{n_0},$$

where s^+ is the solution of $Q_k x = c_k$. Combining (4.6) and the last inequality we obtain

$$\|s^+\|_{Q_k}\,\frac{\sigma}{(\lambda_k\|A\|^2+1)+\sigma\sqrt{\lambda_k\|A\|^2+1}} \;\ge\; \|s_{n_0} - x^+\|_{Q_k}\,.$$

Thus, it follows from Lemma A.1 (see Remark A.2) that

$$\|(\lambda_k A^*A + I)s_{n_0} - \lambda_k A^*(y^\delta - Ax^\delta_{k-1})\| \;\le\; \sigma\|s_{n_0}\|\,.$$

Therefore, r_{n_0} in Algorithm 3 satisfies

$$\|r_{n_0}\| = \|Q_k s_{n_0} - c_k\| = \|(\lambda_k A^*A + I)s_{n_0} - \lambda_k A^*(y^\delta - Ax^\delta_{k-1})\| \le \sigma\|s_{n_0}\|,$$

[1] Here $\lceil x \rceil = \operatorname{ceiling}(x)$ denotes the least integer that is greater than or equal to x.

from where we conclude that $n^\ddagger \leq n_0$. The estimate in (4.5) is a direct consequence of this fact and (4.7). $\square$

We conclude this section verifying Assumption 3.9 for Algorithm 3.

Item 1) From Step [1] of Algorithm 3 follows $T_{0,\lambda_k}(x_{k-1}^\delta, y^\delta) = x_{k-1}^\delta + s_0 = x_{k-1}^\delta$;

Item 2) The continuity of $T_{n,\lambda_k} : X \times Y \to X$ follows from the continuity of the mapping $c_k : X \times Y \ni (x, y) \mapsto c_k(x, y) = \lambda_k A^*(y - Ax) \in X$ and Lemma A.4.

Item 3) From Step [4] of Algorithm 3 and Lemma 4.2 follow the existence of $n_k^\delta = n^\ddagger$, satisfying estimate (4.5), such that $\widetilde{x}_k^\delta = x_{k-1}^\delta + s_{n_k^\delta} = T_{n_k^\delta,\lambda_k}(x_{k-1}^\delta, y^\delta)$ for each $1 \leq k \leq k^*(\delta) + 1$;

Item 4) Let (δ_j), (y^{δ_j}), $(x_k^{\delta_j}) = x_0^{\delta_j}, \ldots, x_{n_{k,j}}^{\delta_j}$ and $(\widetilde{x}_k^{\delta_j}) = \widetilde{x}_1^{\delta_j}, \ldots, \widetilde{x}_{n_{k,j}+1}^{\delta_j}$ be given as in Assumption 3.9; for each $j \in \mathbb{N}$ and $1 \leq k \leq k^*(\delta_j)$ we denote by $n_{k,j} \in \mathbb{N}$ the number of iterations required by Algorithm 3 (with $y^\delta = y^{\delta_j}$, $x_{k-1}^\delta = x_{k-1}^{\delta_j}$, $Q_k = \lambda_k A^* A + I$ and $c_k = \lambda_k A^*(y^{\delta_j} - A x_{k-1}^{\delta_j})$ to reach the stop criteria (see Step [4] in Algorithm 3).
Define $\kappa := \operatorname{cond}(Q_k)$. Since $k \in \mathbb{N}$ is fixed, it follows from (4.5)

$$n_{k,j} \;\leq\; N_k \;:=\; \left\lceil \ln\Big(\frac{(\lambda_k\|A\|^2 + 1) + \sigma\sqrt{\lambda_k\|A\|^2 + 1}}{\sigma/2}\Big) \,\Big[\ln\Big(\frac{\sqrt{\kappa}+1}{\sqrt{\kappa}-1}\Big)\Big]^{-1} \right\rceil$$

for $j = 1, 2, \ldots$ with $k^*(\delta^j) \geq k$, proving item 4 of Assumption 3.9.

5 Numerical experiments

sec:numerics

In what follows the *Inverse Potential Problem* [19, 9, 4, 29] is used to test the numerical efficiency of the ret-iT method. All computations are performed using MATLAB® R2017a, running on an Intel® Core™ i9-10900 CPU.

The 2D Inverse Potential Problem

Let $\Omega \subset \mathbb{R}^2$ be an open bounded domain with Lipschitz-continuous boundary $\partial\Omega$, and assume that $u \in H_0^1(\Omega)$ is the weak solution of the elliptic boundary value problem (BVP)

$$-\Delta u \;=\; x\,,\ \text{in}\ \Omega\,, \quad u \;=\; 0\,,\ \text{on}\ \partial\Omega\,, \tag{5.1}$$

eq:ipp

where $x \in L_2(\Omega)$ is a source function. See [24] for the solution theory of this particular problem.

The corresponding inverse problem is known as Inverse Potential Problem (IPP) [19]. It consists of recovering an L_2–function x, from measurements of the Neumann data of its corresponding potential on the boundary of Ω, i.e. we aim to recover $x \in L_2(\Omega)$ from $y := u_\nu|_{\partial\Omega}$ (the normal derivative of u at the boundary $\partial\Omega$). Generalizations of this linear inverse problem lead to distinct applications, e.g., Inverse Gravimetry [20, 29], EEG [6], and EMG [30].

Given the usual Sobolev space $H^1(\Omega)$, let $H^{1/2}(\partial\Omega)$ be the space of boundary traces of functions in $H^1(\Omega)$, and $H^{-1/2}(\partial\Omega)$ the dual of $H^{1/2}(\partial\Omega)$. The linear direct problem is modeled by the operator $A : L_2(\Omega) \to H^{-1/2}(\partial\Omega)$, where $Ax := u_\nu|_{\partial\Omega}$ and $u \in H_0^1(\Omega)$ is the unique solution of (5.1) Using this notation, the IPP can be modeled in the form (1.2), where the available noisy data $y^\delta \in L_2(\partial\Omega)$ satisfies (1.1).

Let $A^* : H^{1/2}(\partial\Omega) \to L_2(\Omega)$ be the harmonic extension operator, i.e. given $v \in H^{1/2}(\partial\Omega)$, A^*v solves

$$-\Delta(A^*v) = 0\,, \text{ in } \Omega\,, \quad A^*v = v\,, \text{ on } \partial\Omega\,. \tag{5.2}$$

eq:harmoni

Note that A^* is the dual of A since, from the above definitions, given $x \in L^2(\Omega)$ and $v \in H^{1/2}(\partial\Omega)$,

$$\int_{\partial\Omega} Axv = \int_\Omega \nabla u \cdot \nabla A^*v + \int_\Omega xA^*v = \int_\Omega xA^*v,$$

as expected.

We next rewrite (5.1) and (5.2) in a single formulation. Consider $\phi \in H^1(\Omega)$ weak solution of $-\Delta\phi = f$ in Ω and $\phi = g$ on $\partial\Omega$ for given $f \in L^2(\Omega)$ and $g \in H^{1/2}(\partial\Omega)$. The primal-mixed formulation states that $(\phi, \psi) \in H^1(\Omega) \times H^{-1/2}(\partial\Omega)$ is such that [10]

$$\begin{aligned} \int_\Omega \nabla\phi \cdot \nabla v + \int_{\partial\Omega} \psi v &= \int_\Omega fv \quad \text{for all } v \in H^1(\Omega), \\ \int_{\partial\Omega} \mu\phi \quad\quad\quad\quad &= g \quad\quad\;\; \text{for all } \mu \in H^{-1/2}(\partial\Omega). \end{aligned} \tag{5.3}$$

eq:pmf

Above, the "integrals" involving elements of $H^{1/2}(\partial\Omega)$ and $H^{-1/2}(\partial\Omega)$ actually denote the duality between these spaces. Integrating by parts the first equation in (5.3), we see that $-\Delta\phi = f$ and that $\psi = \partial\phi/\partial n$ over the boundary. The second equation in (5.3) imposes the Dirichlet boundary condition $\phi = g$ weakly.

In what follows, we assume that (5.1) is *regular* in the sense that the normal derivative of the solution in $L_2(\partial\Omega)$ and not only in $H^{-1/2}(\partial\Omega)$, i.e. an extra regularity holds [3].

Discretization using finite elements

To discretize the above problems, we use finite element methods as described in [10]. Consider a regular, quasi-uniform triangulation $\mathcal{T}_h$ with elements of characteristic length $h > 0$. Note that $\mathcal{T}_h$ defines a partion on $\partial\Omega$, and we define a new boundary partition Γ_h such that each edge of Γ_h is the union of two edges of $\mathcal{T}_h$.

We define the spaces of piecewise linear and piecewise constant functions

$$\begin{aligned} V_h &= \{v_h \in \mathcal{C}(\overline{\Omega}) : \, v_h|_K \in P_1(K), \, K \in \mathcal{T}_h\}, \\ Q_h &= \{\mu_h \in L^2(\partial\Omega) : \, v_h|_K \in P_0(e), \, e \in \Gamma_h\}, \end{aligned}$$

and search for $(u_h, \psi_h) \in V_h \times Q_h$ such that

$$\begin{aligned} \int_\Omega \nabla u_h \cdot \nabla v_h + \int_{\partial\Omega} \psi v_h &= \int_\Omega xv_h \quad \text{for all } v_h \in V_h, \\ \int_{\partial\Omega} \mu_h u_h \quad\quad\quad\quad &= 0 \quad\quad\quad \text{for all } \mu_h \in Q_h. \end{aligned}$$

Such formulation computes the approximation ψ_h of the normal derivative of the exact solution without post-processing.

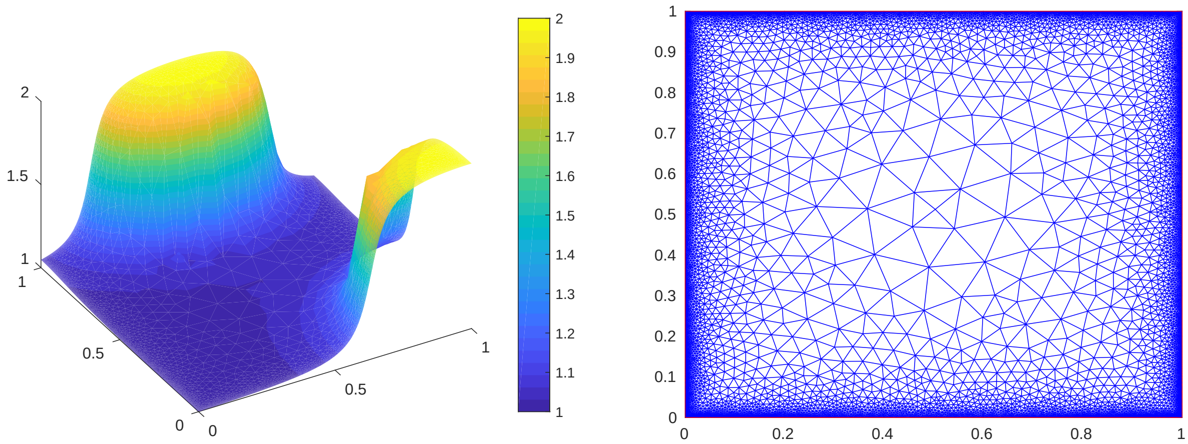

Figure 1: Inverse Potential Problem setup. (LEFT) Ground truth $x^\star$; (RIGHT) Finite element mesh used to solve the inverse problem.

fig:IPP-se

Experiments with noisy data

The numerical tests discussed in this section follow [2, 4] in the experimental setup. Here, $d = 2$ and $\Omega = (0,1) \times (0,1)$. Moreover, the unknown ground truth $x^\star$ is assumed to be an H^1-function with sharp gradients (see Figure 1). The setup of our experiments is as follows:

— Problem (5.1) is solved for $x = x^\star$ and data $y^\star$ are computed.

— We added to data $y^\star$ a normally distributed noise with zero mean and suitable variance for achieving a prescribed relative noise level.

— Two distinct noise scenarios are considered, where the relative noise level $\|y^\star - y^\delta\|/\|y^\star\|$ corresponds to 0.1% and 2%.

— The constant function $x_0 \equiv 1.5$ is used as initial guess.

— We set $\tau = 1.5$, $\sigma = 0.9$ and $\lambda_k = (\frac{3}{2})^k$ in the ret-iT method (Algorithm 2).

— Algorithm 3 is used to compute $\widetilde{x}_k^\delta$ in Step [3.1] of the ret-iT method.

The finite element mesh used to solve the inverse problem (see Figure 1) is coarser than the one used to generate the data $y^\star$. This strategy is adopted in order to avoid inverse crimes [1, 8].

Noisy level of 0.1% The ret-iT method (Algorithm 2) is implemented using the above described setup.

For comparison purposes, the iT method is implemented for solving the IPP (the same experimental setup is used). In order to compute the step of the iT method, see (2.1), the CG method with standard stopping rule is used (i.e. Algorithm 3 with stopping rule $\|r_{i+1}\| > tol$ in Step [3]).[2] The discrepancy principle is used as stopping rule for the iT method, i.e. the iteration stops when $\|Ax^\delta_{\mathrm{iT,k}} - y^\delta\| \le \tau\delta$ for the first time.

The ret-iT reaches the stop criteria after $k^*(\delta) = 11$ steps. The iterate x^δ_{11} as well as the corresponding relative iteration error $|x^\star - x^\delta_{11}|/|x^\star|$ are depicted in Figure 2. The iT method is implemented with $\tau = 1.5$, $\lambda_k = (\frac{3}{2})^k$ and initial guess x_0; it reaches the stop criteria after $k^*(\delta) = 11$ steps.

[2] We choose $tol := 10^{-6}$, which is the default MATLAB tolerance.

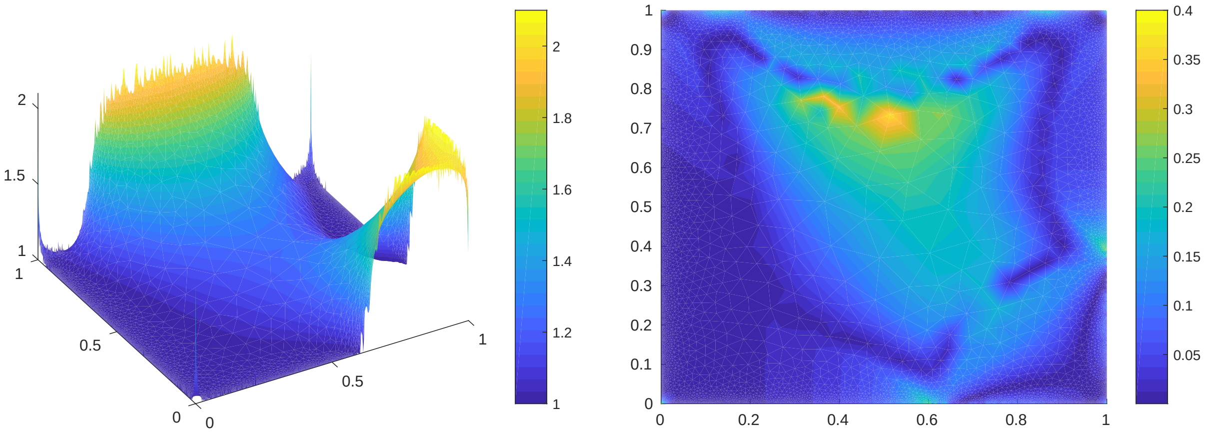

Figure 2: Noise level 0.1%. The ret-iT method reaches the stop criteria after $k^*(\delta) = 11$ steps. (LEFT) Iterate x_{11}^δ; (RIGHT) Relative iteration error $|x^\star - x_{11}^\delta|/|x^\star|$. fig:IPP-no

In Figure 3 we compare the performances of ret-iT and iT methods. In this figure, the evolution of the relative iteration error (TOP) and the evolution of the relative residual (BOTTOM) are plotted for the ret-iT and iT methods (both plots are in logarithmic scale). Additionaly, in order to compare the numerical effort of these two methods, the number of accumulated CG-steps computed in the inner iterations is plotted (CENTER). In Table 1 the number of CG-steps needed in each iteration of the ret-iT method is comapred with the number of CG-steps needed in each iteration of the iT method, for $k = 1, \ldots, 11 = k^*(\delta)$.

	Iteration number										
	1	2	3	4	5	6	7	8	9	10	11
ret-iT method	1	1	1	1	1	1	2	2	1	2	3
iT method	4	4	5	5	6	7	8	8	9	10	12

Table 1: Noise level 0.1%. Number of CG-steps required to compute $\widetilde{x}_k^\delta$ in each iteration of the ret-iT method vs. number of CG-steps required to compute $x_{\mathrm{iT,k}}^\delta$ in the iT method. tab:CG-ste

Noisy level of 2% The ret-iT and iT methods are implemented using the above described setup. Both methods reach the stop criteria after $k^*(\delta) = 7$ steps. In Figure 4 the performances of ret-iT and iT methods are compared. In Table 2 the number of CG-steps needed in each iteration of the ret-iT and iT methods is plotted.

	Iteration number						
	1	2	3	4	5	6	7
ret-iT method	1	1	1	1	1	1	2
iT method	4	4	5	5	6	7	8

Table 2: Noise level 2%. Number of CG-steps in each iteration of ret-iT and iT methods. tab:CG-ste

Comparison with other inexact iT method (relative residual tolerant)

Inexact Newton type methods are a well established alternative for solving nonlinear ill-posed operator equations of the type $F(x) = y^\delta$, e.g., the REGINN iteration [25, 27] or the inexact Levenberg-

Marquardt (LM) method [17]. In both REGINN and inexact LM methods the iterative step $x_k^\delta = x_{k-1}^\delta + s_k$ is computed using updates s_k satisfying $\|F'(x_{k-1}^\delta)s - y^\delta + F(x_{k-1}^\delta)\| < \sigma_k \|F(x_{k-1}^\delta) - y^\delta\|$ for apropriately chosen $\sigma_k < 1$.[3] In particular, in the inexact LM method the update $s_{k,\mathrm{LM}}$ is an aproximate solution of the equation $(\lambda_k F'(x_{k-1}^\delta)^* F'(x_{k-1}^\delta) + I)s = \lambda_k F'(x_{k-1}^\delta)^*(y^\delta - F(x_{k-1}^\delta))$.[4]

In the linear setting (1.2) the inexact LM method above reduces to an inexact iT type method with update $s_{k,\mathrm{iT}}$ satisfying $\|(\lambda_k A^*A + I)s - \lambda_k A^*(y^\delta - Ax_{k-1}^\delta)\| \le \sigma_k \|\lambda_k A^*(y^\delta - Ax_{k-1}^\delta)\|$ (compare with the update of the ret-iT method in (4.1), (4.2) and (4.3)). Such an uptade $s_{k,\mathrm{iT}}$ can be obtained using a slight variation of CG in Algorithm 3, where the stop criterion of the repeat-loop is substituted by: "**if** $\big(\|r_{n+1}\| > \mu_n \|c_k\|\big)$ **then**", with $c_k = \lambda_k A^*(y^\delta - Ax_{k-1})$ and $\mu_n \in (0,1)$ chosen as in [25, Lemma 3.2].

We refer to this inexact iT method as rr-iT, since the calculation of the inexact update $s_{k,\mathrm{iT}}$ is based on relative residual tolerance. In the sequel we revisit the experiment with noise level 0.1% and compare the numerical performance of both inexact iT methods, ret-iT and rr-iT. The discrepancy principle is used as stopping rule for both methods.

In Figure 5 the evolution of residual and iteration error for both inexact iT methods ret-iT and rr-iT are plotted. In Table 3 the number of CG-steps needed in each iterative step of ret-iT and rr-iT is plotted. The rr-iT method requires 12 iterations to reach the stop criterion (one more than ret-iT). Notice that rr-iT requires 37% more CG-steps than ret-iT in order to reach the same stop criterion.

	Iteration number											
	1	2	3	4	5	6	7	8	9	10	11	12
ret-iT method	1	1	1	1	1	1	2	2	1	2	3	
rr-iT method	1	1	2	2	2	2	2	2	2	2	2	2

tab:inexact-iT

Table 3: Noise level 0.1%. Number of CG-steps required in the implementation of the ret-iT method vs. number of CG-steps required in the implementation of the rr-iT method.

6 Conclusions

sec:conclusion

In this notes we propose an inexact iterated-Tikhonov method with relative error tolerance, here called ret-iT, for solving linear ill-posed problems.

The advantage of adopting the relative error tolerant strategy in the computation of the iterative step of the iT method is evident: computationaly, it is less expensive to obtain a solution to the relaxed problem (2.3a) than to calculate the exact solution (up to the computer precision) to the original problem (2.1). In the first numerical experiment (noise level of 0.1%) the ret-iT method required a total of 16 CG-steps to reach the stop criteria, while the iT method required 78 CG-steps (see Table 1 and Figure 3 (CENTER)). In the second experiment (noise level of 2%) the ret-iT method required 8 CG-steps to reach the stop criteria, while the iT method required 39 CG-steps (see Table 2 and Figure 4 (CENTER)). In the third experiment (noise level of 0.1% revisited) the ret-iT method is compared with the rr-iT method (an inexact relative residual tolerant iT method). The rr-iT required a total of 22 CG-steps, 37% more than the ret-iT method (see Table 3 and Figure 5 (CENTER)).

An additional benefit of this strategy lies in the fact that the progress towards the solution of the iterate in (2.3b) is quantitatively "almost as good" as the one obtained in the iT method (2.2); see Lemma 2.4 and the pictures on the (TOP) of Figures 3 and 4.

[3] Clearly, s_k is an approximation for the exact Newton step $s_k^\dagger$, which satisfies $\|F'(x_k)s - y^\delta + F(x_k)\| = 0$.
[4] Here $(\lambda_k) > 0$ is an appropriately chosen sequence of Lagrange multipliers [17].

Appendix (auxiliary results)

In order to prove that the CG method in Algorithm 3 satisfies Assumption 3.9, three auxiliary results are required. The first auxiliary result gives a sufficient condition for $\widetilde{x}$ to satisfy the error criterion in (2.3a), using the norm defined by the quadratic form to be minimized by the CG method.

Lemma A.1. *Given $x_{k-1}^\delta \in X$ and $\lambda_k > 0$; define Q_k, b_k as in (4.1), and $\|\cdot\|_{Q_k}$ as in (4.4). Let x^+ be the solution of $Q_k x = b_k$, i.e.*

$$x^+ \;=\; x_{k-1}^\delta - Q_k^{-1}\lambda_k A^*(Ax_{k-1}^\delta - y^\delta).$$

For $\sigma \geq 0$, if $\widetilde{x} \in X$ and

$$\|\widetilde{x} - x^+\|_{Q_k} \;\leq\; \frac{\sigma}{(\lambda_k\|A\|^2+1) + \sigma\sqrt{\lambda_k\|A\|^2+1}}\,\|x^+ - x_{k-1}^\delta\|_{Q_k}\,, \tag{A.1}$$

then $\|\lambda_k A^*(A\widetilde{x} - y^\delta) + \widetilde{x} - x_{k-1}^\delta\| \leq \sigma\|\widetilde{x} - x_{k-1}^\delta\|$.

Proof. To prove this lemma, it suffices to show that $\beta := \sigma\|\widetilde{x}-x_{k-1}^\delta\| - \|\lambda_k A^*(A\widetilde{x}-y^\delta)+\widetilde{x}-x_{k-1}^\delta\| \geq 0$, whenever $\widetilde{x} \in X$ satisfies (A.1). First observe that, for any $z \in X$

$$\|Q_k z\| \;\leq\; \sqrt{\lambda_k\|A\|^2+1}\,\|z\|_{Q_k}\,,\;\; \|z\| \;\leq\; \|z\|_{Q_k}\,,\;\; \|z\|_{Q_k} \;\leq\; \sqrt{\lambda_k\|A\|^2+1}\,\|z\|\,. \tag{A.2}$$

It follows from the definition of x^+ and the operator Q_k that

$$\lambda_k A^*(A\widetilde{x} - y^\delta) + \widetilde{x} - x_{k-1}^\delta \;=\; Q_k(\widetilde{x} - x^+).$$

Combining this identity with the above definition of β, the triangle inequality, and (A.2), we obtain

$$\begin{aligned}
\beta &= \sigma\|\widetilde{x} - x_{k-1}^\delta\| - \|Q_k(\widetilde{x} - x^+)\| \\
&\geq \sigma(\|x^+ - x_{k-1}^\delta\| - \|\widetilde{x} - x^+\|) - \|Q_k(\widetilde{x} - x^+)\| \\
&\geq \frac{\sigma}{\sqrt{\lambda_k\|A\|^2+1}}\,\|x^+ - x_{k-1}^\delta\|_{Q_k} - \Big[\sigma + \sqrt{\lambda_k\|A\|^2+1}\Big]\|\widetilde{x} - x^+\|_{Q_k}\,.
\end{aligned}$$

Thus, it follows from (A.1) that $\beta \geq 0$, concluding the proof. □

An immediate consequence of Lemma A.1 is the fact that $\widetilde{x}_k^\delta = \widetilde{x}$ satisfies (2.3a), the problem in Step [3.1] of Algorithm 2.

Remark A.2. *In the context of (4.3), Lemma A.1 reads:*
Given $x_{k-1}^\delta \in X$ and $\lambda_k > 0$; define Q_k as in (4.1) and $c_k = \lambda_k A^*(y^\delta - Ax_{k-1}^\delta)$. Let s^+ be the solution of $Q_k x = c_k$, i.e. $s^+ = Q_k^{-1}\lambda_k A^*(y^\delta - Ax_{k-1}^\delta)$. For $\sigma \geq 0$, if $\widetilde{s} \in X$ and

$$\|\widetilde{s} - x^+\|_{Q_k} \;\leq\; \frac{\sigma}{(\lambda_k\|A\|^2+1) + \sigma\sqrt{\lambda_k\|A\|^2+1}}\,\|s^+\|_{Q_k} \tag{A.3}$$

then $\|(\lambda_k A^*A + I)\widetilde{s} - \lambda_k A^*(y^\delta - Ax_{k-1}^\delta)\| \leq \sigma\|\widetilde{s}\|$.

The second auxiliary result in this appendix provides a convergence rate for the CG method. For a proof we refer the reader to [11, 13].

Lemma A.3. *Let Q be a bounded, self-adjoint, coercive operator with condition number κ, let s^+ be* *the solution of $Qx = c$, and let (s_n) be the sequence generated by CG method. Then*

$$\frac{\|s_n - s^+\|_Q}{\|s_0 - s^+\|_Q} \;\leq\; 2\Big[\Big(\frac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1}\Big)^n + \Big(\frac{\sqrt{\kappa}+1}{\sqrt{\kappa}-1}\Big)^n\Big]^{-1} \;\leq\; 2\Big(\frac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1}\Big)^n, \;\; n = 1, 2, \ldots$$

Here $\|\cdot\|_Q$ is defined as in (4.4).

The last auxiliary result in this appendix addresses the continuity of the n^{th} step of the CG method for each fixed $n \in \mathbb{N}$.

lem:CG-c **Lemma A.4.** *Let $\{c^i\}$ be a sequence in X, $c \in X$, and let Q be a self-adjoint, coercive, and bounded linear operator on X. Define, for $n = 0, 1, \ldots$,*

$$\begin{aligned} s_n^i &= \arg\min \tfrac{1}{2}\langle x, Qx\rangle - \langle x, c^i\rangle \quad \textit{for } x \in K_n(Q, c^i)\,, \quad i = 1, 2, \ldots\,; \\ s_n &= \arg\min \tfrac{1}{2}\langle x, Qx\rangle - \langle x, c\rangle \quad \textit{for } x \in K_n(Q, c)\,. \end{aligned}$$

Here $K_n(Q, c)$ are the Krylov spaces generated by Q and c, i.e. $K_0(Q, c) = \{0\}$, $K_n(Q, c) = span(c, \ldots, Q^{n-1}c)$, $n = 1, 2, \ldots$ If

$$c^i \to c \qquad \textit{as} \qquad i \to \infty$$

then

$$s_n^i \to s_n \qquad \textit{as} \qquad i \to \infty \qquad \textit{for} \;\; n = 0, 1, \ldots \tag{A.4}$$

Proof. Fix n. There are $\xi \in \mathbb{R}^n$ and $\xi^i \in \mathbb{R}^n$, for $i = 1, 2, \ldots$, such that

$$s_n = \sum_{j=1}^{n} \xi_j Q^{j-1} c \qquad , \qquad s_n^i = \sum_{j=1}^{n} \xi_j^i Q^{j-1} c^i \tag{A.5}$$ eq:ssi

We shall consider whether $c, \ldots, Q^{n-1}c$ are linearly independent (LI) or not.
a) Suppose first that $c, \ldots, Q^{n-1}c$ are LI.
Under this assumption, ξ^i as in (A.5) is univocally determined

Define $M \in \mathbb{R}^{n\times n}$ and $\eta \in \mathbb{R}^n$ as

$$M_{jk} \;=\; \langle Q^{j-1}c, Q^k c\rangle \;\; 1 \leq j, k \leq n\,; \qquad \eta_j = \langle Q^{j-1}c, c\rangle \;\; 1 \leq j \leq n.$$

Since $Qs_n - c \perp Q^{j-1}c$ for $j = 1, \ldots, n$

$$\sum_{k=1}^{n} M_{jk}\xi_k \;=\; \langle Qs_n, Q^{j-1}c\rangle \;=\; \eta_j \qquad j = 1, \ldots, n$$

Hence $M\xi = \eta$. As $c, \ldots, Q^{-1}c$ are LI, M is non-singular and

$$\xi = M^{-1}\eta.$$

Define $M^i \in \mathbb{R}^{n\times n}$ and $\eta^i \in \mathbb{R}^n$, for $i = 1, 2, \ldots$, as

$$M_{jk}^i \;=\; \langle Q^{j-1}c^i, Q^k c^i\rangle \;\; 1 \leq j, k \leq n\,; \qquad \eta_j^i = \langle Q^{j-1}c^i, c^i\rangle \;\; 1 \leq j \leq n.$$

By the same token, $Qs_n^i - c^i \perp Q^{j-1}c^i$ for $j = 1, \ldots, n$, and $M^i\xi^i = \eta^i$. Since $M^i \to M$, for i large enough M^i is non-singular. Therefore,

$$\xi^i = (M^i)^{-1}\eta^i \qquad\qquad \text{(for } i \text{ large enough).} \tag{A.6}$$ eq:

Since $M^i \to M$ and $\eta^i \to \eta$ as $i \to \infty$,

$$\xi^i \to \xi \quad \text{and} \quad s_n^i = \sum \xi_j^i Q^{j-1} c^i \to \sum \xi_j Q^{j-1} c = s_n$$

as $i \to \infty$.

b) Suppose that $c, \ldots, Q^{n-1}c$ are linearly dependent (LD).

Define

$$\tilde{s}_n^i = \sum_{j=1}^{n} \xi_j Q^{j-1} c^i$$

and let $\varphi^i : X \to X$, for $i = 1, 2, \ldots$, and $\varphi : X \to X$ be

$$\varphi(x) = \tfrac{1}{2}\langle x, Qx\rangle - \langle x, c\rangle, \quad \varphi^i(x) = \tfrac{1}{2}\langle x, Qx\rangle - \langle x, c^i\rangle \tag{A.7}$$

As $\tilde{s}_n^i \in K_n(c^i, Q)$ and s_n^i minimizes φ^i on $K_n(c^i, Q)$

$$\varphi^i(Q^{-1}c^i) + \tfrac{1}{2}\left\| s_n^i - Q^{-1}c^i \right\|_Q^2 = \varphi^i(s_n^i) \le \varphi^i\left(\tilde{s}_n^i\right)$$

Since $c, \ldots, Q^{n-1}$ are LD, s_n is the global minimizer of φ and $s_n = Q^{-1}c$. To end the proof, observe that

$$\tilde{s}_n^i \to s_n, \qquad \varphi^i\left(\tilde{s}_n^i\right) \to \varphi(s_n), \tag{A.8}$$

$$Q^{-1}c^i \to Q^{-1}c = s_n, \qquad \varphi^i\left(Q^{-1}c^i\right) \to \varphi(s_n). \tag{A.9}$$

□

Acknowledgments

AL acknowledges support from the AvH Foundation. ALM acknowledges the financial support funding agency FAPERJ, from the State of Rio de Janeiro. BFS acknowledges support from the agencies CNPq (grants 311300/2020-0, 430868/2018-9) and FAPERJ (grant E-26/203.318/2017).

References

[1] J. Baumeister, *Stable Solution of Inverse Problems*, Advanced Lectures in Mathematics, Friedr. Vieweg & Sohn, Braunschweig, 1987. MR 889048

[2] R. Boiger, A. Leitão, and B.F. Svaiter, *Range-relaxed criteria for choosing the Lagrange multipliers in nonstationary iterated Tikhonov method*, IMA Journal of Numerical Analysis **40** (2020), no. 1, 606–627.

[3] P.G. Ciarlet, *The finite element method for elliptic problems*, Classics in Applied Mathematics, vol. 40, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2002, Reprint of the 1978 original [North-Holland, Amsterdam; MR0520174 (58 #25001)].

[4] A. De Cezaro, A. Leitão, and X.-C. Tai, *On multiple level-set regularization methods for inverse problems*, Inverse Problems **25** (2009), 035004.

[5] A. De Cezaro, J. Baumeister, and A. Leitão, *Modified iterated Tikhonov methods for solving systems of nonlinear ill-posed equations*, Inverse Probl. Imaging **5** (2011), no. 1, 1–17.

[6] A. El Badia and M. Farah, *Identification of dipole sources in an elliptic equation from boundary measurements: application to the inverse EEG problem*, J. Inverse Ill-Posed Probl. **14** (2006), no. 4, 331–353.

[7] H.W. Engl, *On the choice of the regularization parameter for iterated Tikhonov regularization of ill-posed problems*, J. Approx. Theory **49** (1987), no. 1, 55–63.

[8] H.W. Engl, M. Hanke, and A. Neubauer, *Regularization of Inverse Problems*, Kluwer Academic Publishers, Dordrecht, 1996.

[9] F. Frühauf, O. Scherzer, and A. Leitão, *Analysis of Regularization Methods for the Solution of Ill-Posed Problems Involving Discontinuous Operators*, SIAM J. Numer. Anal. **43** (2005), 767–786.

[10] G.N. Gatica, *A simple introduction to the mixed finite element method*, SpringerBriefs in Mathematics, Springer, Cham, 2014, Theory and applications.

[11] G.H. Golub and C.F. Van Loan, *Matrix computations*, fourth ed., Johns Hopkins Studies in the Mathematical Sciences, Johns Hopkins University Press, Baltimore, MD, 2013.

[12] C. W. Groetsch and O. Scherzer, *Non-stationary iterated Tikhonov-Morozov method and third-order differential equations for the evaluation of unbounded operators*, Math. Methods Appl. Sci. **23** (2000), no. 15, 1287–1300.

[13] W. Hackbusch, *Iterative Lösung großer schwachbesetzter Gleichungssysteme*, Leitfäden der Angewandten Mathematik und Mechanik [Guides to Applied Mathematics and Mechanics], vol. 69, B. G. Teubner, Stuttgart, 1991, Teubner Studienbücher Mathematik. [Teubner Mathematical Textbooks].

[14] M. Hanke, *Conjugate gradient type methods for ill-posed problems*, Longman Scientific & Technical, 1995.

[15] M. Hanke and C. W. Groetsch, *Nonstationary Iterated Tikhonov Regularization*, J. Optim. Theory Appl. **98** (1998), no. 1, 37–53.

[16] M. Hanke, A. Neubauer, and O. Scherzer, *A convergence analysis of Landweber iteration for nonlinear ill-posed problems*, Numer. Math. **72** (1995), 21–37.

[17] Martin Hanke, *A regularizing Levenberg-Marquardt scheme, with applications to inverse groundwater filtration problems*, Inverse Problems **13** (1997), no. 1, 79–95.

[18] ——— , *Regularizing properties of a truncated Newton-CG algorithm for nonlinear inverse problems*, Numer. Funct. Anal. Optim. **18** (1997), no. 9-10, 971–993.

[19] F. Hettlich and W. Rundell, *Iterative methods for the reconstruction of an inverse potential problem*, Inverse Problems **12** (1996), no. 3, 251–266.

[20] Victor Isakov, *Inverse Problems for Partial Differential Equations*, second ed., Applied Mathematical Sciences, vol. 127, Springer, New York, 2006.

[21] B. Kaltenbacher, A. Neubauer, and O. Scherzer, *Iterative Regularization Methods for Nonlinear Ill-Posed Problems*, Radon Series on Computational and Applied Mathematics, vol. 6, Walter de Gruyter GmbH & Co. KG, Berlin, 2008.

[22] S. Kindermann and A. Neubauer, *On the convergence of the quasioptimality criterion for (iterated) Tikhonov regularization*, Inverse Probl. Imaging **2** (2008), no. 2, 291–299.

[23] A. Kirsch, *An Introduction to the Mathematical Theory of Inverse Problems*, Applied Mathematical Sciences, vol. 120, Springer-Verlag, New York, 1996.

[24] J.L. Lions and E. Magenes, *Non-homogeneous boundary value problems and applications*, vol. 1, Springer, New York, 1972.

[25] A. Rieder, *On the regularization of nonlinear ill-posed problems via inexact Newton iterations*, Inverse Problems **15** (1999), no. 1, 309–327.

[26] ———, *Keine Probleme mit inversen Problemen*, Friedr. Vieweg & Sohn, Braunschweig, 2003, Eine Einführung in ihre stabile Lösung. [An introduction to their stable solution].

[27] ———, *Inexact Newton regularization using conjugate gradients as inner iteration*, SIAM J. Numer. Anal. **43** (2005), no. 2, 604–622.

[28] O. Scherzer, *Convergence rates of iterated Tikhonov regularized solutions of nonlinear ill-posed problems*, Numer. Math. **66** (1993), no. 2, 259–279.

[29] K. van den Doel, U. M. Ascher, and A. Leitão, *Multiple Level Sets for Piecewise Constant Surface Reconstruction in Highly Ill-Posed Problems*, Journal of Scientific Computing **43** (2010), no. 1, 44–66.

[30] Kees van den Doel, Uri M. Ascher, and Dinesh K. Pai, *Computed myography: three-dimensional reconstruction of motor functions from surface EMG data*, Inverse Problems **24** (2008), no. 6, 065010, 17.

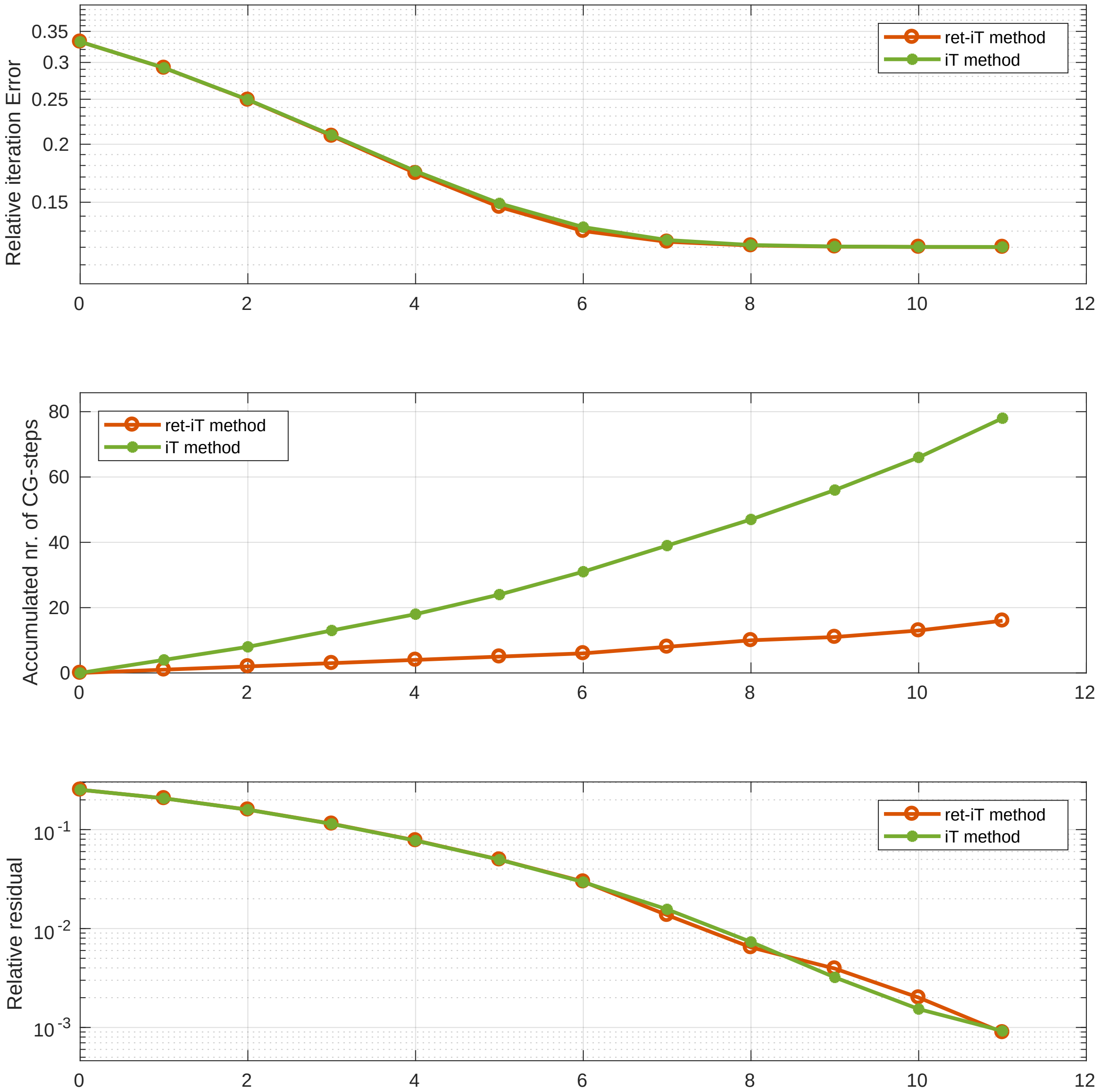

Figure 3: Noise level 0.1%. Comparison betwen ret-iT method and iT method. (TOP) Relative iteration error; (CENTER) Accumulated number of CG-steps computed in the inner iterations; (BOTTOM) Relative residual. fig:IPP-ev

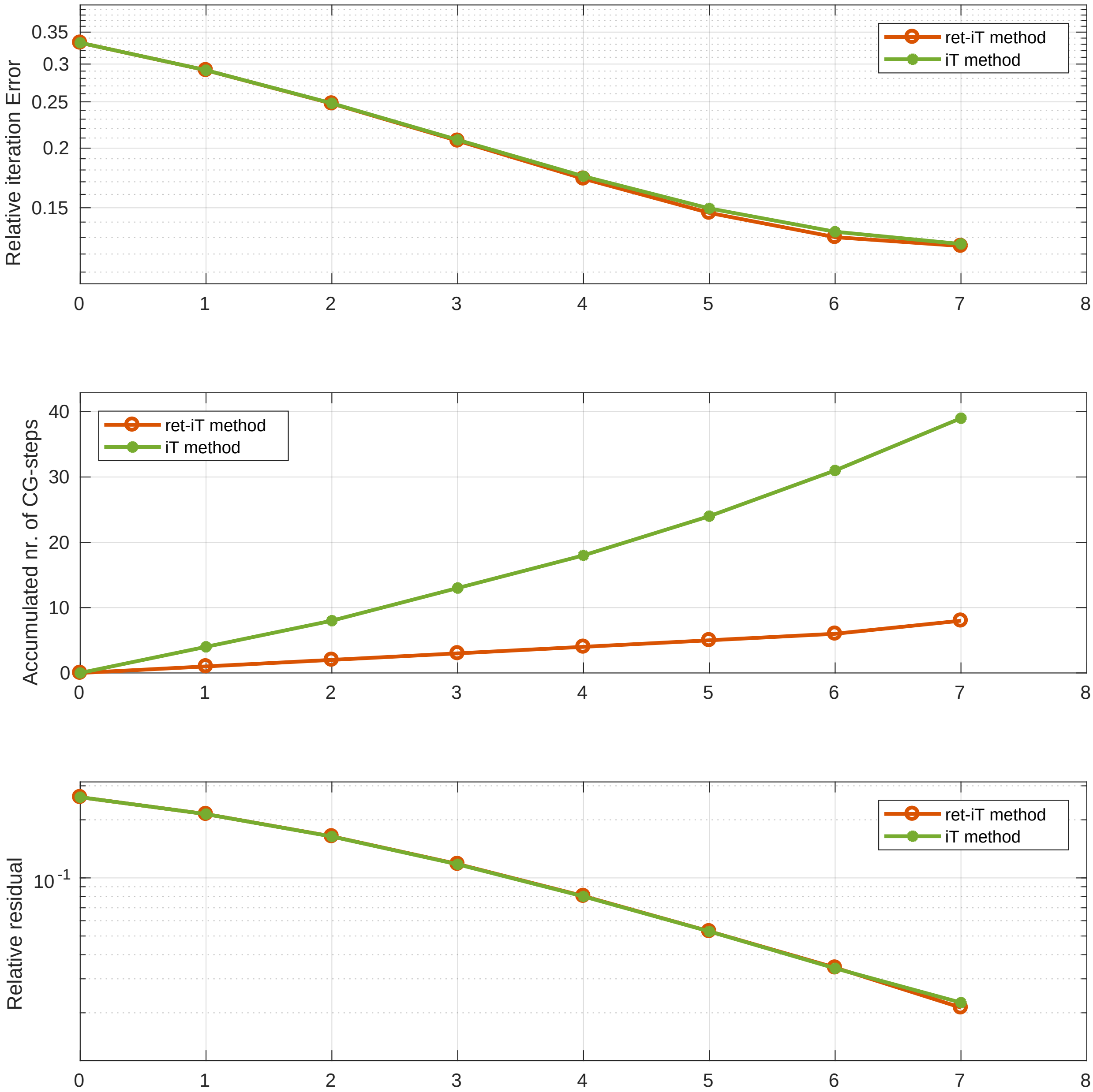

Figure 4: Noise level 2%. Comparison betwen ret-iT method and iT method. (TOP) Relative iteration error; (CENTER) Accumulated number of CG-steps computed in the inner iterations; (BOTTOM) Relative residual.

`fig:IPP-e`

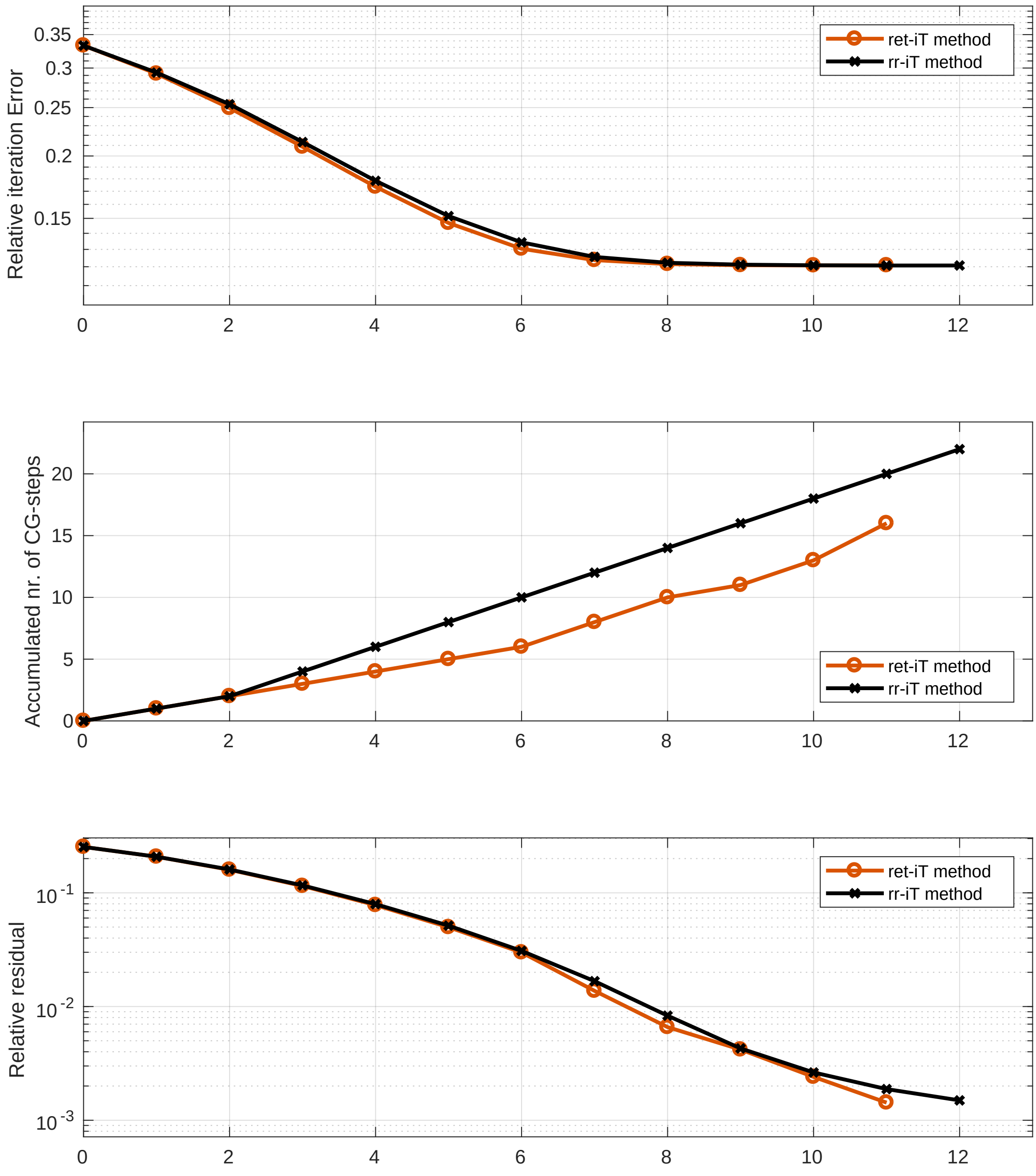

Figure 5: Noise level 0.1% revisited. Comparison betwen inexact iT methods ret-iT and rr-iT. (TOP) Relative iteration error; (CENTER) Accumulated number of CG-steps computed in the inner iterations; (BOTTOM) Relative residual.

fig:inexac