Iterated fractional Tikhonov regularization

Davide Bianchi[*] Alessandro Buccini[*] Marco Donatelli[*]
Stefano Serra-Capizzano[*]

Abstract

Fractional Tikhonov regularization methods have been recently proposed to reduce the oversmoothing property of the Tikhonov regularization in standard form, in order to preserve the details of the approximated solution. Their regularization and convergence properties have been previously investigated showing that they are of optimal order. This paper provides saturation and converse results on their convergence rates. Using the same iterative refinement strategy of iterated Tikhonov regularization, new iterated fractional Tikhonov regularization methods are introduced. We show that these iterated methods are of optimal order and overcome the previous saturation results. Furthermore, nonstationary iterated fractional Tikhonov regularization methods are investigated, establishing their convergence rate under general conditions on the iteration parameters. Numerical results confirm the effectiveness of the proposed regularization iterations.

1 Introduction

We consider linear operator equations of the form

$$Kx = y\,, \tag{1.1}$$

where $K : \mathcal{X} \to \mathcal{Y}$ is a compact linear operator between Hilbert spaces $\mathcal{X}$ and $\mathcal{Y}$. We assume y to be attainable, i.e., that problem (1.1) has a solution $x^\dagger = K^\dagger y$ of minimal norm. Here $K^\dagger$ denotes the (Moore-Penrose) generalized inverse operator of K, which is unbounded when K is compact, with infinite dimensional range. Hence problem (1.1) is ill-posed and has to be regularized in order to compute a numerical solution; see [15].

We want to approximate the solution $x^\dagger$ of the equation (1.1), when only an approximation y^δ of y is available with

$$\|y^\delta - y\| \le \delta, \tag{1.2}$$

where δ is called the noise level. Since $K^\dagger y^\delta$ is not a good approximation of $x^\dagger$, we approximate $x^\dagger$ with $x_\alpha^\delta := R_\alpha y^\delta$ where $\{R_\alpha\}$ is a family of continuous operators depending on a parameter α that will be defined later. A classical example is the Tikhonov regularization defined by $R_\alpha = (K^*K + \alpha I)^{-1}K^*$, where I denotes the identity and K^* the adjoint of K, cf. [16].

Using the singular values expansion of K, filter based regularization methods are defined in terms of filters of the singular values, cf. Proposition 3. This is a useful tool for the analysis

[*]Dipartimento di Scienza e Alta Tecnologia, Università dell'Insubria, 22100 Como, Italy

of regularization techniques [3], both for direct and iterative regularization methods [4, 12]. Furthermore, new regularization methods can be defined investigating new classes of filters. For instance, one of the contributes in [6] is the proposal and the analysis of the fractional Tikhonov method. The authors obtain a new class of filtering regularization methods adding an exponent, depending on a parameter, to the filter of the standard Tikhonov method. They provide a detailed analysis of the filtering properties and the optimality order of the method in terms of such further parameter. A different generalization of the Tikhonov method has been recently proposed in [9] with a detailed filtering analysis. Both generalizations are called "fractional Tikhonov regularization" in the literature and they are compared in [7], where the optimality order of the method in [9] is provided as well. To distinguish the two proposals in [6] and [9], we will refer in the following as "fractional Tikhonov regularization" and "weighted Tikhonov regularization", respectively. These variants of the Tikhonov method have been introduced to compute good approximations of non-smooth solutions, since it is well known that the Tikhonov method provides over-smoothed solutions.

In this paper, we firstly provide a saturation result similar to the well-known saturation result for Tikhonov regularization [15]: let $R(K)$ be the range of K and let Q be the orthogonal projector onto $\overline{R(K)}$, if

$$\sup\left\{\|x_\alpha^\delta - x^\dagger\| : \|Q(y-y^\delta)\| \leq \delta\right\} = o(\delta^{\frac{2}{3}}),$$

then $x^\dagger = 0$, as long as $\overline{R(K)}$ is not closed. Such result motivated us to introduce the iterated version of fractional and weighted Tikhonov in the same spirit of the iterated Tikhonov method. We prove that this iterated methods can overcome the previous saturation results. Afterwards, inspired by the works [1, 8] we introduce the nonstationary variants of our iterated methods. Differently from the nonstationary iterated Tikhonov, we have two nonstationary sequences of parameters. In the noise free case, we give sufficient conditions on these sequences to guarantee the convergence providing also the corresponding convergence rates. In the noise case, we show the stability of the proposed iterative schemes proving that they are regularization methods. Finally, few selected examples confirm the previous theoretical analysis, showing that a proper choice of the nonstationary sequence of parameters can provide better restorations compared to the classical iterated Tikhonov with a geometric sequence of regularizzation parameters according to [8].

The paper is organized as follows. Section 2 recalls the basic definition of filter based regularization methods and of optimal order of a regularization method. Fractional Tikhonov regularization with optimal order and converse results are studied in Section 3. Section 4 is devoted to saturation results for both variants of fractional Tikhonov regularization. New iterated fractional Tikhonov regularization methods are introduced in Section 5, where the analysis of their convergence rate shows that their are able to overcome the previous saturation results. A nonstationary iterated weighted Tikhonov regularization is investigated in detail in Section 6, while a similar nonstationary iterated fractional Tikhonov regularization is discussed in Section 7. Finally, some numerical examples are reported in Section 8.

2 Preliminaries

As described in the Introduction, we consider a compact linear operator $K : \mathcal{X} \to \mathcal{Y}$ between Hilbert spaces $\mathcal{X}$ and $\mathcal{Y}$ (over the field $\mathbb{R}$ or $\mathbb{C}$) with given inner products $\langle \cdot, \cdot \rangle_{\mathcal{X}}$ and $\langle \cdot, \cdot \rangle_{\mathcal{Y}}$, respectively. Hereafter we will omit the subscript for the inner product as it will be clear in the context. If $K^* : \mathcal{Y} \to \mathcal{X}$ denotes the adjoint of K (i.e., $\langle Kx, y \rangle = \langle x, K^* y \rangle$), then we indicate with $(\sigma_n; v_n, u_n)_{n\in\mathbb{N}}$ the singular value expansion (s.v.e.) of K, where $\{v_n\}_{n\in\mathbb{N}}$ and $\{u_n\}_{n\in\mathbb{N}}$ are a complete orthonormal system of eigenvectors for K^*K and KK^*, respectively, and $\sigma_n > 0$ are written in decreasing order, with 0 being the only accumulating point for the sequence $\{\sigma_n\}_{n\in\mathbb{N}}$. If $\mathcal{X}$ is not finite dimensional, then $0 \in \sigma(K^*K)$, the spectrum of K^*K, namely $\sigma(K^*K) = \{0\} \cup \bigcup_{n=1}^{\infty} \{\sigma_n^2\}$. Finally, $\sigma(K)$ denotes the closure of $\bigcup_{n=1}^{\infty} \{\sigma_n\}$, i.e., $\sigma(K) = \{0\} \cup \bigcup_{n=1}^{\infty} \{\sigma_n\}$.

Let now $\{E_{\sigma^2}\}_{\sigma^2 \in \sigma(K^*K)}$ be the spectral decomposition of the self-adjoint operator K^*K. Then from well-known facts from functional analysis [14] we can write $f(K^*K) := \int f(\sigma^2) dE_{\sigma^2}$, where $f : \sigma(K^*K) \subset \mathbb{R} \to \mathbb{C}$ is a bounded Borel measurable function and $\langle E x_1, x_2 \rangle$ is a regular complex Borel measure for every $x_1, x_2 \in \mathcal{X}$. The following equalities hold

$$Kx = \sum_{m=1}^{+\infty} \sigma_m \langle x, v_m \rangle u_m, \qquad x \in \mathcal{X}, \tag{2.1}$$

$$K^* y = \sum_{m=1}^{+\infty} \sigma_m \langle y, u_m \rangle v_m, \qquad y \in \mathcal{Y}, \tag{2.2}$$

$$f(K^*K)x := \int_{\sigma(K^*K)} f(\sigma^2) dE_{\sigma^2} x = \sum_{m=1}^{\infty} f(\sigma_m^2) \langle x, v_m \rangle v_m, \tag{2.3}$$

$$\langle f(K^*K)x_1, x_2 \rangle = \int_{\sigma(K^*K)} f(\sigma^2) d\langle E_{\sigma^2} x_1, x_2 \rangle = \sum_{m=1}^{\infty} f(\sigma_m^2) \overline{\langle y, v_m \rangle} \langle x, v_m \rangle, \tag{2.4}$$

$$\| f(K^*K) \| \le \sup\{ |f(\sigma^2)| : \sigma^2 \in \sigma(K^*K) \}, \tag{2.5}$$

where the series (2.1) and (2.2) converge in the L^2 norms induced by the scalar products of $\mathcal{X}$ and $\mathcal{Y}$, respectively. If f is a continuous function on $\sigma(K^*K)$ then equality holds in (2.5).

Definition 1 *We define the* generalized inverse $K^\dagger : \mathcal{D}(K^\dagger) \subseteq \mathcal{Y} \to \mathcal{X}$ *of a compact linear operator* $K : \mathcal{X} \to \mathcal{Y}$ *as*

$$K^\dagger y = \sum_{m:\, \sigma_m > 0} \sigma_m^{-1} \langle y, u_m \rangle v_m, \qquad y \in \mathcal{D}(K^\dagger), \tag{2.6}$$

where

$$\mathcal{D}(K^\dagger) = \left\{ y \in \mathcal{Y} : \sum_{m:\, \sigma_m > 0} \sigma_m^{-2} |\langle y, u_m \rangle|^2 < \infty \right\}.$$

With respect to problem (1.1), we consider the case where only an approximation y^δ of y satisfying the condition (1.2) is available. Therefore $x^\dagger = K^\dagger y$, $y \in \mathcal{D}(K^\dagger)$, cannot be approximated by $K^\dagger y^\delta$, due to the unboundedness of $K^\dagger$, and hence in practice the problem (1.1) is approximated by a family of neighbouring well-posed problems [15].

Definition 2 *By a* regularization method *for* $K^\dagger$ *we call any family of operators*

$$\{R_\alpha\}_{\alpha\in(0,\alpha_0)} : \mathcal{Y} \to \mathcal{X}, \qquad \alpha_0 \in (0,+\infty],$$

with the following properties:

(i) $R_\alpha : \mathcal{Y} \to \mathcal{X}$ *is a bounded operator for every* α.

(ii) For every $y \in \mathcal{D}(K^\dagger)$ *there exists a mapping (rule choice)* $\alpha : \mathbb{R}_+ \times \mathcal{Y} \to (0,\alpha_0) \in \mathbb{R}$, $\alpha = \alpha(\delta, y^\delta)$, *such that*

$$\lim_{\delta\to 0} \sup \left\{ \alpha(\delta, y^\delta) : y^\delta \in \mathcal{Y}, \|y - y^\delta\| \leq \delta \right\} = 0,$$

and

$$\lim_{\delta\to 0} \sup \left\{ \|R_{\alpha(\delta,y^\delta)} y^\delta - K^\dagger y\| : y^\delta \in \mathcal{Y}, \|y - y^\delta\| \leq \delta \right\} = 0.$$

Throughout this paper c is a constant which can change from one instance to the next. For the sake of clarity, if more than one constant will appear in the same line or equation we will distinguish them by means of a subscript.

Proposition 3 *Let* $K : \mathcal{X} \to \mathcal{Y}$ *be a compact linear operator and* $K^\dagger$ *its generalized inverse. Let* $R_\alpha : \mathcal{Y} \to \mathcal{X}$ *be a family of operators defined for every* $\alpha \in (0,\alpha_0)$ *as*

$$R_\alpha y := \sum_{m:\,\sigma_m>0} F_\alpha(\sigma_m)\sigma_m^{-1}\langle y, u_m\rangle v_m, \tag{2.7}$$

where $F_\alpha : [0,\sigma_1] \supset \sigma(K) \to \mathbb{R}$ *is a Borel function such that*

$$\sup_{m:\,\sigma_m>0} |F_\alpha(\sigma_m)\sigma_m^{-1}| = c(\alpha) < \infty, \tag{2.8a}$$

$$|F_\alpha(\sigma_m)| \leq c < \infty, \qquad \text{where } c \text{ does not depend on } (\alpha, m), \tag{2.8b}$$

$$\lim_{\alpha\to 0} F_\alpha(\sigma_m) = 1 \ \text{ point-wise in } \sigma_m. \tag{2.8c}$$

Then R_α *is a regularization method, with* $\|R_\alpha\| = c(\alpha)$, *and it is called* filter based regularization method.

Proof. See [10] and [15]. □

For the sake of notational brevity, we fix the following notation

$$x_\alpha := R_\alpha y, \qquad y \in \mathcal{D}(K^\dagger), \tag{2.9}$$

$$x_\alpha^\delta := R_\alpha y^\delta, \qquad y^\delta \in \mathcal{Y}. \tag{2.10}$$

We report hereafter the definition of optimal order, under the same a-priori assumption given in [15].

Definition 4 *For every given $\nu, \rho > 0$, let*

$$\mathcal{X}_{\nu,\rho} := \left\{ x \in \mathcal{X} : \exists\, \omega \in \mathcal{X}, \|\omega\| \leq \rho, x = (K^*K)^{\frac{\nu}{2}}\omega \right\} \subset \mathcal{X}.$$

A regularization method R_α is called of optimal order *under the a-priori assumption $x^\dagger \in \mathcal{X}_{\nu,\rho}$ if*

$$\Delta(\delta, \mathcal{X}_{\nu,\rho}, R_\alpha) \leq c \cdot \delta^{\frac{\nu}{\nu+1}} \rho^{\frac{1}{\nu+1}}, \tag{2.11}$$

where for any general set $M \subseteq X$, $\delta > 0$ and for a regularization method R_α, we define

$$\Delta(\delta, M, R_\alpha) := \sup \left\{ \|x^\dagger - x_\alpha^\delta\| : x^\dagger \in M,\ \|y - y^\delta\| \leq \delta \right\}.$$

If ρ is not known, as it will be usually the case, then we relax the definition introducing the set

$$\mathcal{X}_\nu := \bigcup_{\rho > 0} \mathcal{X}_{\nu,\rho}$$

and saying that a regularization method R_α is called of optimal order *under the a-priori assumption $x^\dagger \in \mathcal{X}_\nu$ if*

$$\Delta(\delta, \mathcal{X}_\nu, R_\alpha) \leq c \cdot \delta^{\frac{\nu}{\nu+1}}. \tag{2.12}$$

Remark 5 *Since we are concerned with the rate that $\|x^\dagger - x_\alpha^\delta\|$ converges to zero as $\delta \to 0$, the a-priori assumption $x^\dagger \in \mathcal{X}_\nu$ is usually sufficient for the optimal order analysis, requiring that* (2.12) *is satisfied.*

Hereafter we cite a theorem which states sufficient conditions for order optimality, when filtering methods are employed, see [10, Proposition 3.4.3, pag. 58].

Theorem 6 *[10] Let $K : \mathcal{X} \to \mathcal{Y}$ be a compact linear operator, ν and $\rho > 0$, and let $R_\alpha : \mathcal{Y} \to \mathcal{X}$ be a filter based regularization method. If there exists a fixed $\beta > 0$ such that*

$$\sup_{0 < \sigma \leq \sigma_1} |F_\alpha(\sigma)\sigma^{-1}| \leq c \cdot \alpha^{-\beta}, \tag{2.13a}$$

$$\sup_{0 \leq \sigma \leq \sigma_1} |(1 - F_\alpha(\sigma))\sigma^\nu| \leq c_\nu \cdot \alpha^{\beta\nu}, \tag{2.13b}$$

then R_α is of optimal order, under the a-priori assumption $x^\dagger \in \mathcal{X}_{\nu,\rho}$, with the choice rule

$$\alpha = \alpha(\delta, \rho) = O\left(\frac{\delta}{\rho}\right)^{\frac{1}{\beta(\nu+1)}}.$$

If we are concerned just about the rate of convergence with respect to only δ, the preceding theorem can be applied under the a-priori assumption $x^\dagger \in X_\nu$, fitting the proof to the latter case without any effort. On the contrary, below we present a converse result.

Theorem 7 *Let K be a compact linear operator with infinite dimensional range and let R_α be a filter based regularization method with filter function $F_\alpha : [0, \sigma_1] \supset \sigma(K) \to \mathbb{R}$. If there exist ν and $\beta > 0$ such that*

$$(1 - F_\alpha(\sigma))\,\sigma^\nu \geq c\alpha^{\beta\nu} \qquad \text{for } \sigma \in [c'\alpha^\beta, \sigma_1] \tag{2.14}$$

and

$$\|x^\dagger - x_\alpha\| = O(\alpha^{\beta\nu}), \tag{2.15}$$

then $x^\dagger \in \mathcal{X}_\nu$.

Proof. By (2.6) and (2.7), it holds

$$\begin{aligned}
\|x^\dagger - x_\alpha\|^2 &= \sum_{\sigma_m>0} (1 - F_\alpha(\sigma_m))^2\,\sigma_m^{-2}|\langle y, u_m\rangle|^2 \\
&= \sum_{\sigma_m>0} (1 - F_\alpha(\sigma_m))^2\,|\langle x^\dagger, v_m\rangle|^2 \\
&= \sum_{\sigma_m>0} [(1 - F_\alpha(\sigma_m))\,\sigma_m^\nu]^2\,\sigma_m^{-2\nu}|\langle x^\dagger, v_m\rangle|^2 \\
&\geq \left(c\alpha^{\beta\nu}\right)^2 \sum_{\sigma_m \geq c'\alpha^\beta} \sigma_m^{-2\nu}|\langle x^\dagger, v_m\rangle|^2.
\end{aligned}$$

thanks to the assumption (2.14). From (2.15) we deduce that

$$\lim_{\alpha^\beta \to 0} \sum_{\sigma_m \geq c'\alpha^\beta} \sigma_m^{-2\nu}|\langle x^\dagger, v_m\rangle|^2 < +\infty.$$

Finally, if we define $\omega := \sum_{\sigma_m>0} \sigma^{-\nu}\langle x^\dagger, v_m\rangle v_m$, then ω is well defined and $(K^*K)^{\nu/2}\,\omega = x^\dagger$, i.e., $x^\dagger \in X_\nu$. □

3 Fractional variants of Tikhonov regularization

In this section we discuss two recent types of regularization methods that generalize the classical Tikhonov method and that were first introduced and studied in [9] and [6].

3.1 Weighted Tikhonov regularization

Definition 8 ([9]) *We call* Weighted Tikhonov *method the filter based method*

$$R_{\alpha,r}y := \sum_{m:\,\sigma_m>0} F_{\alpha,r}(\sigma_m)\sigma_m^{-1}\langle y, u_m\rangle v_m,$$

where the filter function is

$$F_{\alpha,r}(\sigma) = \frac{\sigma^{r+1}}{\sigma^{r+1} + \alpha}, \tag{3.1}$$

for $\alpha > 0$ and $r \geq 0$.

According to (2.9) and (2.10), we fix the following notation

$$x_{\alpha,r} := R_{\alpha,r}y, \qquad y \in \mathcal{D}(K^\dagger), \tag{3.2}$$

$$x^\delta_{\alpha,r} := R_{\alpha,r}y^\delta, \qquad y^\delta \in \mathcal{Y}. \tag{3.3}$$

Remark 9 *The Weighted Tikhonov method can also be defined as the unique minimizer of the following functional,*

$$R_{\alpha,r}y := \operatorname{argmin}_{x\in X} \left\{\|Kx - y\|_W + \alpha\|x\|\right\}, \tag{3.4}$$

where the semi-norm $\|\cdot\|_W$ *is induced by the operator* $W := (KK^*)^{\frac{r-1}{2}}$*. For* $0 \le r < 1$*,* W *is to be intended as the Moore-Penrose (pseudo) inverse. Developing the calculations, it follows that*

$$R_{\alpha,r}y = \left[(K^*K)^{\frac{r+1}{2}} + \alpha I\right]^{-1} (K^*K)^{\frac{r-1}{2}} K^*y. \tag{3.5}$$

That is the reason that motivated us to rename the original method of Hochstenbach and Reichel, that appeared in [9], into weighted Tikhonov method*. In this way it would be easier to distinguish from the* fractional Tikhonov method *introduced by Klann and Ramlau in [6].*

The optimal order of the weighted Tikhonov regularization was proved in [7]. The following proposition restates such result, putting in evidence the dependence on r of ν, and provides a converse result.

Proposition 10 *Let* K *be a compact linear operator with infinite dimensional range. For every given* $r \ge 0$ *the weighted Tikhonov method,* $R_{\alpha,r}$*, is a regularization method of optimal order, under the a-priori assumption* $x^\dagger \in X_{\nu,\rho}$*, with* $0 < \nu \le r+1$*. The best possible rate of convergence with respect to* δ *is* $\|x^\dagger - x^\delta_{\alpha,r}\| = O\left(\delta^{\frac{r+1}{r+2}}\right)$*, that is obtained for* $\alpha = \left(\frac{\delta}{\rho}\right)^{\frac{r+1}{\nu+1}}$ *with* $\nu = r+1$*. On the other hand, if* $\|x^\dagger - x_{\alpha,r}\| = O(\alpha)$ *then* $x^\dagger \in \mathcal{X}_{r+1}$*.*

Proof. For weighted Tikhonov the left-hand side of condition (2.13a) becomes

$$\sup_{0<\sigma\le\sigma_1} \left|\frac{\sigma^r}{\sigma^{r+1}+\alpha}\right|.$$

By derivation, if $r > 0$ then it is straightforward to see that the quantity above is bounded by $\alpha^{-\beta}$, with $\beta = 1/(r+1)$. Similarly, the left-hand side of condition (2.13b) takes the form

$$\sup_{0\le\sigma\le\sigma_1} \left|\frac{\alpha\sigma^\nu}{\sigma^{r+1}+\alpha}\right|,$$

and it is easy to check that it is bounded by $\alpha^{\beta\nu}$ if and only if $0 < \nu \le r+1$. From Theorem 6, as long as $0 < \nu \le r+1$, with $r > 0$, if $x^\dagger \in \mathcal{X}_{\nu,\rho}$ then we find order optimality (2.11) and the best possible rate of convergence obtainable with respect to δ is $O(\delta^{\frac{r+1}{\nu+1}})$, for $\nu = r+1$.

On the contrary, with $\beta = 1/(r+1)$ and $\nu = r+1$, we deduce that

$$|(1 - F_{\alpha,r}(\sigma))\,\sigma^\nu| = \frac{\alpha\sigma^\nu}{\sigma^{r+1}+\alpha} \ge \frac{1}{2}\alpha, \qquad \text{for } \sigma \in [\alpha^\beta, \sigma_1].$$

Therefore, if $\|x^\dagger - x_{\alpha,r}\| = O(\alpha)$ then $x^\dagger \in \mathcal{X}_\nu$ by Theorem 7. □

3.2 Fractional Tikhonov regularization

Here we introduce the *fractional Tikhonov* method defined and discussed in [6].

Definition 11 ([6]) *We call* Fractional Tikhonov *method the filter based method*

$$R_{\alpha,\gamma}y := \sum_{m:\,\sigma_m>0} F_{\alpha,\gamma}(\sigma_m)\sigma_m^{-1}\langle y, u_m\rangle v_m,$$

where the filter function is

$$F_{\alpha,\gamma}(\sigma) = \frac{\sigma^{2\gamma}}{(\sigma^2+\alpha)^{\gamma}}, \tag{3.6}$$

for $\alpha > 0$ and $\gamma \geq 1/2$.

Note that $F_{\alpha,\gamma}$ is well-defined also for $0 < \gamma < 1/2$, but the condition (2.8a) requires $\gamma \geq 1/2$ to guarantee that $F_{\alpha,\gamma}$ is a filter function.

We use the notation for $x_{\alpha,\gamma}$ and $x^{\delta}_{\alpha,\gamma}$ like in equations (3.2) and (3.3), respectively. The optimal order of the fractional Tikhonov regularization was proved in [6, Proposition 3.2]. The following proposition restates such result including also $\gamma = 1/2$ and provides a converse result.

Proposition 12 *The extended fractional Tikhonov filter method is a regularization method of optimal order, under the a-priori assumption $x^{\dagger} \in X_{\nu,\rho}$, for every $\gamma \geq 1/2$ and $0 < \nu \leq 2$. The best possible rate of convergence with respect to δ is $\|x^{\dagger} - x^{\delta}_{\alpha,\gamma}\| = O\left(\delta^{\frac{2}{3}}\right)$, that is obtained for $\alpha = \left(\frac{\delta}{\rho}\right)^{\frac{2}{\nu+1}}$ with $\nu = 2$. On the other hand, if $\|x^{\dagger} - x_{\alpha,\gamma}\| = O(\alpha)$ then $x^{\dagger} \in \mathcal{X}_2$.*

Proof. Condition (2.8a) is verified for $\gamma \geq 1/2$ and the same holds for conditions (2.8b) and (2.8c). Deriving the filter function, it is immediate to see that equation (2.13a) is verified for $\gamma \geq 1/2$, with $\beta = 1/2$. It remains to check equation (2.13b):

$$\begin{aligned}
(1 - F_{\alpha,\gamma}(\sigma))\,\sigma^{\nu} &= \frac{(\sigma^2+\alpha)^{\gamma} - \sigma^{2\gamma}}{(\sigma^2+\alpha)^{\gamma}}\sigma^{\nu} \\
&= \frac{\left(\frac{\sigma^2}{\alpha}+1\right)^{\gamma} - \left(\frac{\sigma^2}{\alpha}\right)^{\gamma}}{\left(\frac{\sigma^2}{\alpha}+1\right)^{\gamma-1}} \cdot \frac{\alpha\sigma^{\nu}}{\sigma^2+\alpha} \\
&= h\left(\frac{\sigma^2}{\alpha}\right) \cdot (1 - F_{\alpha,1}(\sigma))\,\sigma^{\nu},
\end{aligned}$$

where $h(x) = \frac{(x+1)^{\gamma} - x^{\gamma}}{(x+1)^{\gamma-1}}$ is monotone, $h(0) = 1$ for every γ, and $\lim_{x\to\infty} h(x) = \gamma$. Namely $h(x) \in (\gamma, 1]$ for $0 \leq \gamma \leq 1$ and $h(x) \in [1, \gamma)$ for $\gamma \geq 1$. Therefore we deduce that

$$\gamma\,(1 - F_{\alpha,1}(\sigma)) \leq (1 - F_{\alpha,\gamma}(\sigma)) \leq (1 - F_{\alpha,1}(\sigma)), \qquad \text{for } 0 \leq \gamma \leq 1, \tag{3.7}$$

$$(1 - F_{\alpha,1}(\sigma)) \leq (1 - F_{\alpha,\gamma}(\sigma)) \leq \gamma\,(1 - F_{\alpha,1}(\sigma)), \qquad \text{for } \gamma \geq 1, \tag{3.8}$$

from which we infer that

$$\sup_{\sigma\in[0,\sigma_1]} |(1-F_{\alpha,\gamma}(\sigma))\,\sigma^\nu| \le \max\{1,\gamma\} \sup_{\sigma\in[0,\sigma_1]} |(1-F_{\alpha,1}(\sigma))\,\sigma^\nu| \le c\alpha^{\frac{\nu}{2}}, \tag{3.9}$$

since $F_{\alpha,1}(\sigma)$ is standard Tikhonov, that is of optimal order, with $\beta = 1/2$ and for every $0 < \nu \le 2$, see [15]. On the contrary, with $\beta = 1/2$ and $\nu = 2$, and by equations (3.7) and (3.8), we deduce that

$$(1-F_{\alpha,\gamma}(\sigma))\,\sigma^2 \ge \min\{1,\gamma\}\,(1-F_{\alpha,1}(\sigma))\,\sigma^2 \ge \frac{1}{2}\alpha, \qquad \text{for } \sigma \in [\alpha^{\frac{1}{2}}, \sigma_1]. \tag{3.10}$$

Therefore, if $\|x^\dagger - x_{\alpha,r}\| = O(\alpha)$ then $x^\dagger \in \mathcal{X}_2$ by Theorem 7. □

4 Saturation results

The following proposition deals with a saturation result similar to a well known result for classic Tikhonov, cf. [15, Proposition 5.3].

Proposition 13 (Saturation for weighted Tikhonov regularization) *Let $K : \mathcal{X} \to \mathcal{Y}$ be a compact linear operator with infinite dimensional range and $R_{\alpha,r}$ be the corresponding family of weighted Tikhonov regularization operators in Definition 8. Let $\alpha = \alpha(\delta, y^\delta)$ be any parameter choice rule. If*

$$\sup\left\{\|x^\delta_{\alpha,r} - x^\dagger\| : \|Q(y-y^\delta)\| \le \delta\right\} = o(\delta^{\frac{r+1}{r+2}}), \tag{4.1}$$

then $x^\dagger = 0$, where we indicated with Q the orthogonal projector onto $\overline{R(K)}$.

Proof. Define

$$\begin{aligned} &\delta_m := \sigma_m^{r+2}, && y^\delta_m := y + \delta_m u_m \text{ so that } \|y - y^\delta_m\| \le \delta_m, \\ &\alpha_m := \alpha(\delta_m, y^\delta_m), && x_m := x_{\alpha_m,r}, \qquad x^\delta_m := x^{\delta_m}_{\alpha_m,r}. \end{aligned}$$

By the assumption that K has not finite dimensional range, then $\sigma_m > 0$ for every m and $\lim_{m\to\infty} \sigma_m = 0$. According to Remark 9, from equation (3.5) we have

$$x^\delta_m - x^\dagger = R_{\alpha_m,r} y^\delta_m - x^\dagger = R_{\alpha_m,r} y + \delta_m R_{\alpha_m,r} u_m - x^\dagger = x_m - x^\dagger + \delta_m F_{\alpha_m,r}(\sigma_m)\sigma_m^{-1} v_m$$

and hence by (3.1)

$$\|x^\delta_m - x^\dagger\|^2 = \|x_m - x^\dagger\|^2 + 2\frac{\delta_m \sigma_m^r}{\sigma_m^{r+1} + \alpha_m}\langle x_m - x^\dagger, v_m\rangle + \left(\frac{\delta_m \sigma_m^r}{\sigma_m^{r+1} + \alpha_m}\right)^2.$$

From the choice of $\delta_m := \sigma_m^{r+2}$ follows that

$$\begin{aligned} \left(\delta_m^{-\frac{r+1}{r+2}}\|x^\delta_m - x^\dagger\|\right)^2 &\ge \frac{2}{\delta_m^{\frac{r+1}{r+2}} + \alpha_m}\langle x_m - x^\dagger, v_m\rangle + \left(\frac{\delta_m^{\frac{r+1}{r+2}}}{\delta_m^{\frac{r+1}{r+2}} + \alpha_m}\right)^2 \\ &= \frac{2}{1 + \delta_m^{-\frac{r+1}{r+2}}\alpha_m}\delta_m^{-\frac{r+1}{r+2}}\langle x_m - x^\dagger, v_m\rangle + \left(\frac{1}{1 + \delta_m^{-\frac{r+1}{r+2}}\alpha_m}\right)^2. \end{aligned} \tag{4.2}$$

By (3.5),

$$\left((K^*K)^{\frac{r+1}{2}} + \alpha_m I\right)(x^\dagger - x_m^\delta) = (K^*K)^{\frac{r+1}{2}} x^\dagger + \alpha_m x^\dagger - (K^*K)^{\frac{r-1}{2}} K^* y_m^\delta$$
$$= \alpha_m x^\dagger - \delta_m (K^*K)^{\frac{r-1}{2}} K^* u_m, \tag{4.3}$$

so that

$$\alpha_m \|x^\dagger\| = O(\delta_m + \|x^\dagger - x_m^\delta\|). \tag{4.4}$$

Since, by assumption, $\|x^\dagger - x_m^\delta\| = o(\delta_m^{\frac{r+1}{r+2}})$, it follows from (4.4) that if $x^\dagger \neq 0$, then

$$\lim_{m\to\infty} \alpha_m \delta_m^{-\frac{r+1}{r+2}} = 0. \tag{4.5}$$

Now, by (4.1) and (4.5) applied to inequality (4.2) it follows that $0 \geq 1$,which is a contradiction. Hence $x^\dagger = 0$. □

Note that for $r = 1$ (classical Tikhonov) the previous proposition gives exactly Proposition 5.3 in [15]. On the other hand, taking a large r, it is possible to overcome the saturation result of classical Tikhonov obtaining a convergence rate arbitrary close to $O(\delta)$.

A similar saturation result can be proved also for the fractional Tikhonov regularization in Definition 11.

Proposition 14 (Saturation for fractional Tikhonov regularization) *Let $K : \mathcal{X} \to \mathcal{Y}$ be a compact linear operator with infinite dimensional range and let $R_{\alpha,\gamma}$ be the corresponding family of fractional Tikhonov regularization operators in Definition 11, with fixed $\gamma \geq 1/2$. Let $\alpha = \alpha(\delta, y^\delta)$ be any parameter choice rule. If*

$$\sup\left\{\|x_{\alpha,\gamma}^\delta - x^\dagger\| : \|Q(y - y^\delta)\| \leq \delta\right\} = o(\delta^{\frac{2}{3}}), \tag{4.6}$$

then $x^\dagger = 0$, where we indicated with Q the orthogonal projector onto $\overline{R(K)}$.

Proof. If $\gamma = 1$, the thesis follows from the saturation result for standard Tikhonov [15, Proposition 5.3]. For $\gamma \neq 1$, recalling that

$$x_{\alpha,\gamma} - x^\dagger = \sum_{\sigma_m > 0} (F_{\alpha,\gamma}(\sigma_m) - 1)\, \sigma_m^{-1} \langle y, u_m \rangle v_m,$$

by equations (3.7) and (3.8), we obtain

$$\|x_{\alpha,\gamma} - x^\dagger\| > c\|x_{\alpha,1} - x^\dagger\|, \tag{4.7}$$

where $c = \min\{1, \gamma\}$ and $x_{\alpha,1}$ is standard Tikhonov. Let us define

$$\phi_\gamma(y) := \|x_{\alpha,\gamma} - x^\dagger\|.$$

Then, by the continuity of ϕ_γ, there exists $\delta > 0$ such that, for every $y^\delta \in \overline{B}_\delta(y)$, we find

$$\phi_\gamma(y^\delta) > c \cdot \phi_1(y^\delta),$$

with $\overline{B}_\delta(y)$ being the closure of the ball of center y and radius δ. Passing to the sup we obtain that

$$\sup\left\{\|x^\delta_{\alpha,\gamma} - x^\dagger\| : \|Q(y - y^\delta)\| \leq \delta\right\} \geq c \cdot \sup\left\{\|x^\delta_{\alpha,1} - x^\dagger\| : \|Q(y - y^\delta)\| \leq \delta\right\}. \tag{4.8}$$

Therefore, using relation (4.6), we deduce

$$\sup\left\{\|x^\delta_{\alpha,1} - x^\dagger\| : \|y - y^\delta\| \leq \delta\right\} = o(\delta^{\frac{2}{3}}), \tag{4.9}$$

and the thesis follows again from the saturation result for standard Tikhonov, cf. [15, Proposition 5.3]. □

Differently from the weighted Tikhonov regularization, for the fractional Tikhonov method, it is not possible to overcome the saturation result of classical Tikhonov, even for a large γ.

5 Stationary iterated regularization

We define new iterated regularization methods based on weighed and fractional Tikhonov regularization using the same iterative refinement strategy of iterated Tikhonov regularization [1,15]. We will show that the iterated methods go beyond the saturation results proved in the previous section. In this section the regularization parameter will still be α with the iteration step, n, assumed to be fixed. On the contrary, in Section 6, we will analyze the nonstationary counterpart of this iterative method, in which α will be replaced by a pre-fixed sequence $\{\alpha_n\}$ and we will be concerned on the rate of convergence with respect to the index n.

5.1 Iterated weighted Tikhonov regularization

We propose now an iterated regularization method based on weighted Tikhonov

Definition 15 (Stationary iterated weighted Tikhonov) *We define the* stationary iterated weighted Tikhonov method *(*SIWT*) as*

$$\begin{cases} x^0_{\alpha,r} := 0; \\ \left((K^*K)^{\frac{r+1}{2}} + \alpha I\right) x^n_{\alpha,r} := (K^*K)^{\frac{r-1}{2}} K^* y + \alpha x^{n-1}_{\alpha,r}, \end{cases} \tag{5.1}$$

with $\alpha > 0$ and $r \geq 0$, or equivalently

$$\begin{cases} x^0_{\alpha,r} := 0 \\ x^n_{\alpha,r} := \operatorname{argmin}_{x\in X} \left\{\|Kx - y\|_W + \alpha\|x - x^{n-1}_{\alpha,r}\|\right\}, \end{cases} \tag{5.2}$$

where $\|\cdot\|_W$ is the semi-norm introduced in (3.4)*. We define $x^{n,\delta}_{\alpha,r}$ as the n-th iteration of weighted Tikhonov if $y = y^\delta$.*

Proposition 16 *For any given $n \in \mathbb{N}$ and $r > 0$, the SIWT in (5.1) is a filter based regularization method, with filter function*

$$F^{(n)}_{\alpha,r}(\sigma) = \frac{(\sigma^{r+1}+\alpha)^n - \alpha^n}{(\sigma^{r+1}+\alpha)^n}. \tag{5.3}$$

Moreover, the method is of optimal order, under the a-priori assumption $x^\dagger \in X_{\nu,\rho}$, for $r > 0$ and $0 < \nu \leq n(r+1)$, with best convergence rate $\|x^\dagger - x^{n,\delta}_{\alpha,r}\| = O(\delta^{\frac{n(r+1)}{1+n(r+1)}})$, that is obtained for $\alpha = (\frac{\delta}{\rho})^{\frac{n(r+1)}{1+\nu}}$, with $\nu = n(r+1)$. On the other hand, if $\|x^\dagger - x^n_{\alpha,r}\| = O(\alpha^n)$, then $x^\dagger \in X_{n(r+1)}$.

Proof. Multiplying both sides of (5.1) by $\left((K^*K)^{\frac{r+1}{2}} + \alpha I\right)^{n-1}$ and iterating the process, we get

$$\begin{aligned}\left((K^*K)^{\frac{r+1}{2}} + \alpha I\right)^n x^n_{\alpha,r} &= \left\{\sum_{j=0}^{n-1} \alpha^j \left((K^*K)^{\frac{r+1}{2}} + \alpha I\right)^{n-1-j}\right\} (K^*K)^{\frac{r-1}{2}} K^* y \\ &= \left[\left((K^*K)^{\frac{r+1}{2}} + \alpha I\right)^n - \alpha^n I\right] (K^*K)^{-1} K^* y.\end{aligned}$$

Therefore, the filter function in (2.7) is equal to

$$F^{(n)}_{\alpha,r}(\sigma) = \frac{(\sigma^{r+1}+\alpha)^n - \alpha^n}{(\sigma^{r+1}+\alpha)^n},$$

as we stated. Condition (2.8c) is straightforward to verify. Moreover, note that

$$\begin{aligned}F^{(n)}_{\alpha,r}(\sigma) &= \frac{(\sigma^{r+1}+\alpha)^n - \alpha^n}{(\sigma^{r+1}+\alpha)^n} \\ &= \frac{\sigma^{r+1}}{\sigma^{r+1}+\alpha} \cdot \frac{\left(\sum_{j=0}^{n-1} \alpha^j (\sigma^{r+1}+\alpha)^{n-1-j}\right)}{(\sigma^{r+1}+\alpha)^{n-1}} \\ &= F_{\alpha,r}(\sigma) \cdot \left(1 + \left(\frac{\alpha}{\sigma^{r+1}+\alpha}\right) + \cdots + \left(\frac{\alpha}{\sigma^{r+1}+\alpha}\right)^{n-1}\right),\end{aligned}$$

from which it follows that

$$F_{\alpha,r}(\sigma) \leq F^{(n)}_{\alpha,r}(\sigma) \leq n F_{\alpha,r}(\sigma). \tag{5.4}$$

Therefore, conditions (2.8a), (2.8b) and (2.13a) follows immediately by the regularity of the weighted Tikhonov filter method for $r > 0$ and by the order optimality for $r > 0$. Finally, condition (2.13b) becomes

$$\sup_{\sigma \in [0,\sigma_1]} \left| \frac{\alpha^n \sigma^\nu}{(\sigma^{r+1}+\alpha)^n} \right|,$$

and deriving one checks that it is bounded by $\alpha^{\beta\nu}$, with $\beta = 1/(r+1)$, if and only if $0 < \nu \leq n(r+1)$. Applying now Proposition 6 the rest of the thesis follows.

On the contrary, if we define $\beta = 1/(r+1)$ and $\nu = n(r+1)$, then we deduce that

$$\left(1 - F^{(n)}_{\alpha,r}(\sigma)\right)\sigma^\nu = \frac{\alpha^n \sigma^\nu}{(\sigma^{r+1}+\alpha)^n} \geq \frac{1}{2^n}\alpha^n \qquad \text{for } \sigma \in [\alpha^\beta, \sigma_1].$$

Therefore, if $\|x^\dagger - x^n_{\alpha,r}\| = O(\alpha^n)$, then by Theorem 7 it follows that $x^\dagger \in X_{n(r+1)}$. $\square$

If n is large, then we note that the convergence rate approaches $O(\delta)$ also for a fixed small r. The study of the convergence for increasing n and fixed α will be dealt with in Section 6.

5.2 Iterated fractional Tikhonov regularization

With the same path as in the previous subsection, we propose here the stationary iterated version of the fractional Tikhonov method.

Definition 17 (Stationary iterated fractional Tikhonov) *We define the* stationary iterated fractional Tikhonov method *(*SIFT*) as*

$$\begin{cases} x^0_{\alpha,\gamma} := 0; \\ (K^*K+\alpha I)^\gamma\, x^n_{\alpha,\gamma} := (K^*K)^{\gamma-1}K^*y + [(K^*K+\alpha I)^\gamma - (K^*K)^\gamma]\, x^{n-1}_{\alpha,\gamma}, \end{cases} \tag{5.5}$$

with $\gamma \geq 1/2$. We define $x^{n,\delta}_{\alpha,\gamma}$ for the n-th iteration of fractional Tikhonov if $y = y^\delta$.

Proposition 18 *For any given $n \in \mathbb{N}$ and $\gamma \geq 1/2$, the SIFT in* (5.5) *is a filter based regularization method, with filter function*

$$F^{(n)}_{\alpha,\gamma}(\sigma) = \frac{(\sigma^2+\alpha)^{\gamma n} - \left[(\sigma^2+\alpha)^\gamma - \sigma^{2\gamma}\right]^n}{(\sigma^2+\alpha)^{\gamma n}}. \tag{5.6}$$

Moreover, the method is of optimal order, under the a-priori assumption $x^\dagger \in X_{\nu,\rho}$, for $\gamma \geq 1/2$ and $0 < \nu \leq 2n$, with best convergence rate $\|x^\dagger - x^{n,\delta}_{\alpha,\gamma}\| = O(\delta^{\frac{2n}{2n+1}})$, that is obtained for $\alpha = (\frac{\delta}{\rho})^{\frac{2n}{\nu+1}}$, with $\nu = 2n$. On the other hand, if $\|x^\dagger - x^n_{\alpha,\gamma}\| = O(\alpha^n)$, then $x^\dagger \in X_{2n}$.

Proof. Multiplying both sides of (5.6) by $(K^*K+\alpha I)^{(n-1)\gamma}$ and iterating the process, we get

$$\begin{aligned} (K^*K+\alpha I)^{n\gamma}\, x^n_{\alpha,\gamma} &= \left\{ \sum_{j=0}^{n-1} (K^*K+\alpha I)^{j\gamma}\, [(K^*K+\alpha I)^\gamma - (K^*K)^\gamma]^{n-1-j} \right\} (K^*K)^{\gamma-1}K^*y \\ &= \{(K^*K+\alpha I)^{\gamma n} - [(K^*K+\alpha I)^\gamma - (K^*K)^\gamma]^n\}\, (K^*K)^{-1}K^*y, \end{aligned}$$

where we used the fact that $(K^*K+\alpha I)^{-\gamma}$ and $[(K^*K+\alpha I)^\gamma - (K^*K)^\gamma]$ commute. Therefore, the filter function in (2.7) is given by

$$F^n_{\alpha,\gamma}(\sigma) = \frac{(\sigma^2+\alpha)^{\gamma n} - \left[(\sigma^2+\alpha)^\gamma - \sigma^{2\gamma}\right]^n}{(\sigma^2+\alpha)^{\gamma n}},$$

as we stated. We observe that

$$
\begin{aligned}
F_{\alpha,\gamma}^{(n)}(\sigma) &= \frac{(\sigma^2+\alpha)^{\gamma n} - \left[\left(\sigma^2+\alpha\right)^{\gamma} - \sigma^{2\gamma}\right]^n}{(\sigma^2+\alpha)^{\gamma n}} \\
&= \frac{\sigma^{2\gamma}}{(\sigma^2+\alpha)^{\gamma}} \cdot \frac{1}{(\sigma^2+\alpha)^{\gamma(n-1)}} \cdot \sum_{j=0}^{n-1} (\sigma^2+\alpha)^{\gamma j} \left[(\sigma^2+\alpha)^{\gamma} - \sigma^{2\gamma}\right]^{n-1-j} \\
&= \frac{\sigma^{2\gamma}}{(\sigma^2+\alpha)^{\gamma}} \cdot \left\{ 1 + \left[1 - \left(\frac{\sigma^2}{\sigma^2+\alpha}\right)^{\gamma}\right] + \cdots + \left[1 - \left(\frac{\sigma^2}{\sigma^2+\alpha}\right)^{\gamma}\right]^{n-1} \right\},
\end{aligned}
$$

from which we deduce that

$$
F_{\alpha,\gamma}^{(n)}(\sigma) \le n F_{\alpha,\gamma}(\sigma). \tag{5.7}
$$

Therefore, since $F_{\alpha,\gamma}$ is a regularization method of optimal order, conditions (2.8a), (2.8b) and (2.13a) are satisfied. Moreover, it is easy to check condition (2.8c) and so we get the regularity for the method. It remains to check condition (2.13b) for the order optimality.

From equations (3.7) and (3.8) we deduce that

$$
\begin{aligned}
1 - F_{\alpha,\gamma}^{(n)}(\sigma) &= \left[\frac{(\sigma^2+\alpha)^{\gamma} - \sigma^{2\gamma}}{(\sigma^2+\alpha)^{\gamma}}\right]^n \\
&= \left[1 - \frac{\sigma^{2\gamma}}{(\sigma^2+\alpha)^{\gamma}}\right]^n \\
&= (1 - F_{\alpha,\gamma}(\sigma))^n \\
&\le (\max\{1,\gamma\})^n \,(1 - F_{\alpha,1}(\sigma))^n \\
&= c\left(1 - F_{\alpha,1}^n(\sigma)\right),
\end{aligned} \tag{5.8}
$$

where $F_{\alpha,1}(\sigma)$ is the standard Tikhonov filter and $F_{\alpha,1}^{(n)}(\sigma)$ is the filter function of the stationary iterated Tikhonov, i.e., $F_{\alpha,1}^{(n)}(\sigma) = \frac{(\sigma^2+\alpha)^n - \alpha^n}{(\sigma^2+\alpha)^n}$. Now condition (2.13b) follows from the properties of stationary iterated Tikhonov, with $\beta = 1/2$ and $0 < \nu \le 2n$, see [12, p. 124]. By applying Proposition 6 we get the best convergence rate, $O(\delta^{\frac{2n}{2n+1}})$.

On the contrary, set $\beta = 1/2$ and $\nu = 2n$. First, let us observe that from equations (5.8) and (3.7), (3.8), we infer that

$$
1 - F_{\alpha,\gamma}^{(n)}(\sigma) \ge (\min\{1,\gamma\})^n \left(1 - F_{\alpha,1}^{(n)}(\sigma)\right).
$$

Then, we deduce that

$$
\begin{aligned}
\left(1 - F_{\alpha,\gamma}^{(n)}(\sigma)\right)\sigma^{\nu} &\ge c\frac{\alpha^n \sigma^{2n}}{(\sigma^2+\alpha)^n} \\
&\ge c\alpha^n \qquad \text{for } \sigma \in [\alpha^{\beta}, \sigma_1].
\end{aligned}
$$

Therefore, if $\|x^{\dagger} - x_{\alpha,\gamma}^n\| = O(\alpha^n)$, then $x^{\dagger} \in X_{2n}$ by Theorem 7. □

The previous proposition shows that, similarly to SIWT, a large n allows to overcome the saturation result in Proposition 14. The study of the convergence for increasing n and fixed α will be dealt with in Section 7.

6 Nonstationary iterated weighted Tikhonov regularization

We introduce a nonstationary version of the iteration (5.1). We study the convergence and we prove that the new iteration is a regularization method.

Definition 19 *Let $\{\alpha_n\}_{n\in\mathbb{N}}, \{r_n\}_{n\in\mathbb{N}} \subset \mathbb{R}_{>0}$ be sequences of positive real numbers. We define a nonstationary iterated weighted Tikhonov method (NSIWT) as follows*

$$\begin{cases} x^0_{\alpha_0,r_0} := 0, \\ \left[(K^*K)^{\frac{r_n+1}{2}} + \alpha_n I\right] x^n_{\alpha_n,r_n} := (K^*K)^{\frac{r_n-1}{2}} K^*y + \alpha_n x^{n-1}_{\alpha_{n-1},r_{n-1}}, \end{cases} \tag{6.1}$$

or equivalently

$$\begin{cases} x^0_{\alpha_0,r_0} := 0, \\ x^n_{\alpha_n,r_n} := \ \operatorname{argmin}_{x\in X} \left\{ \|Kx - y\|_{W_n} + \alpha_n \|x - x^{n-1}_{\alpha_{n-1},r_{n-1}}\| \right\}, \end{cases} \tag{6.2}$$

where $\|\cdot\|_{W_n}$ is the semi-norm introduced by the operator $W_n := (KK^*)^{\frac{r_n-1}{2}}$ and depending on n, due to the non stationary character of r_n.

6.1 Convergence analysis

We are concerned about the properties of the sequence $\{\alpha_n\}$ such that the iteration (6.1) shall converge. To this aim we need some preliminary lemmas, whose proof can be found in the appendix.

Remark 20 *Hereafter, without loss of generality, we will consider $\sigma_1 = 1$, namely $\|K\| = 1$.*

Lemma 21 *Let $\{t_n\}_{n\in\mathbb{N}}$ be a sequence of real numbers such that $0 \le t_n < 1$ for every n. Then*

$$\prod_{n=1}^{\infty}(1 - t_n) > 0 \qquad \text{if and only if} \qquad \sum_{n=1}^{\infty} t_n < \infty. \tag{6.3}$$

Proof. See [13, Theorem 15.5] □

Lemma 22 *Let $\{t_k\}_{k\in\mathbb{N}}$ be a sequence of positive real numbers and let $N > 0$. Then*

$$\sum_{k=1}^{n} t_k \sim c \sum_{k=N}^{n} t_k,$$

with $c > 0$ (in particular, $c = 1$ when $\sum_{k=N}^{\infty} t_k = \sum_{k=1}^{\infty} t_k = \infty$).

Lemma 23 *For every $\lambda \in (0, \infty)$ and for every sequence $\{t_k\}_{k\in\mathbb{N}} \subset (0, \infty)$ such that $\lim_{k\to\infty} t_k = t \in (0, \infty]$, we find*

$$\sum_{k=1}^{n} \frac{1}{t_k} \sim c_\lambda \sum_{k=1}^{n} \frac{\lambda}{\lambda + t_k}, \qquad c_\lambda > 0,$$

where $\sim$ denotes the asymptotic equivalence.

We can now prove a necessary and sufficient condition on the sequence $\{\alpha_n\}$ to have the convergence of NSIWT.

Theorem 24 *The* NSIWT *method* (6.1) *converges to* $x^\dagger \in \mathcal{X}$ *as* $n \to \infty$ *if and only if* $\sum_{k=1}^{n} \frac{\sigma^{r_k+1}}{\sigma^{r_k+1}+\alpha_k}$ *diverges for every* $\sigma > 0$.

Proof. Rewriting equation (6.1) and reminding that $y = Kx^\dagger$, we have

$$\begin{aligned} x^n_{\alpha_n,r_n} &= \left[(K^*K)^{\frac{r_n+1}{2}} + \alpha_n I\right]^{-1} (K^*K)^{\frac{r_n+1}{2}} x^\dagger + \alpha_n \left[(K^*K)^{\frac{r_n+1}{2}} + \alpha_n I\right]^{-1} x^{n-1}_{\alpha_{n-1},r_{n-1}} \\ &= \left\{ I - \alpha_n \left[(K^*K)^{\frac{r_n+1}{2}} + \alpha_n I\right]^{-1} \right\} x^\dagger + \alpha_n \left[(K^*K)^{\frac{r_n+1}{2}} + \alpha_n I\right]^{-1} x^{n-1}_{\alpha_{n-1},r_{n-1}}, \end{aligned}$$

from which it follows that

$$\begin{aligned} x^\dagger - x^n_{\alpha_n,r_n} &= \alpha_n \left[(K^*K)^{\frac{r_n+1}{2}} + \alpha_n I\right]^{-1} (x^\dagger - x^{n-1}_{\alpha_{n-1},r_{n-1}}) \\ &= (\cdots) \text{ iterating the process } n-1 \text{ times} \\ &= \prod_{k=1}^{n} \alpha_k \left[(K^*K)^{\frac{r_k+1}{2}} + \alpha_k I\right]^{-1} x^\dagger \end{aligned} \tag{6.4}$$

since $x^0_{\alpha_0,r_0} := 0$. As a consequence, the method shall converge if and only if

$$\lim_{n\to\infty} \left\| \prod_{k=1}^{n} \alpha_k \left[(K^*K)^{\frac{r_k+1}{2}} + \alpha_k I\right]^{-1} x^\dagger \right\| = 0 \tag{6.5}$$

for every $x^\dagger \in X$, namely, if and only if

$$\lim_{n\to\infty} \int_{\sigma(K^*K)} \left| \prod_{k=1}^{n} \frac{\alpha_k}{\sigma^{r_k+1} + \alpha_k} \right|^2 d\langle E_{\sigma^2} x^\dagger, x^\dagger \rangle = 0 \tag{6.6}$$

for every Borel-measure $\langle Ex^\dagger, x^\dagger \rangle$ induced by $x^\dagger \in X$. Since

$$\left| \prod_{k=1}^{n} \frac{\alpha_k}{\sigma^{r_k+1} + \alpha_k} \right|^2 \leq 1$$

for every n, and since

$$\int_{\sigma(K^*K)} d\langle E_{\sigma^2} x^\dagger, x^\dagger \rangle = \|x^\dagger\|^2,$$

the Dominated Convergence Theorem [13, Theorem 1.34, pag. 26] implies

$$\lim_{n\to\infty} \int_{\sigma(K^*K)} \left| \prod_{k=1}^{n} \frac{\alpha_k}{\sigma^{r_k+1} + \alpha_k} \right|^2 d\langle E_{\sigma^2} x^\dagger, x^\dagger \rangle = \int_{\sigma(K^*K)} \lim_{n\to\infty} \left| \prod_{k=1}^{n} \frac{\alpha_k}{\sigma^{r_k+1} + \alpha_k} \right|^2 d\langle E_{\sigma^2} x^\dagger, x^\dagger \rangle. \tag{6.7}$$

Hence, the NSIWT method is convergent if and only if

$$\prod_{k=1}^{\infty} \frac{\alpha_k}{\sigma^{r_k+1}+\alpha_k} = \prod_{k=1}^{\infty}\left(1-\frac{\sigma^{r_k+1}}{\sigma^{r_k+1}+\alpha_k}\right) = 0, \tag{6.8}$$

for $\langle Ex^\dagger, x^\dagger\rangle$-a.e. σ^2, i.e., for every $\sigma \in \sigma(K)\setminus\{0\}$. Applying now Lemma 21 the thesis follows. □

Corollary 25

(1) If $\sup_{k\in\mathbb{N}}\{r_k\} = r \in [0,\infty)$, then the NSIWT method converges if and only if $\sum_{k=1}^{n}\alpha_k^{-1}$ diverges.

(2) Let $\lim_{k\to\infty} r_k = \infty$ monotonically. If $\left(\sum_{k=1}^{n}\alpha_k^{-1}\right)^{-1} = o(\sigma^{r_n+1})$ for every $\sigma\in\sigma(K)\setminus\{0\}$, then the NSIWT method converges.

Proof. (1) For every $\sigma\in\sigma(K)\setminus\{0\}$, we observe that

$$\sum_{k=1}^{\infty}\frac{\sigma^{r+1}}{\sigma^{r+1}+\alpha_k} \le \sum_{k=1}^{\infty}\frac{\sigma^{r_k+1}}{\sigma^{r_k+1}+\alpha_k} \le \sum_{k=1}^{\infty}\frac{1}{1+\alpha_k} \le \sum_{k=1}^{\infty}\frac{1}{\alpha_k}. \tag{6.9}$$

If the NSIWT method converges then, by Theorem 24 and by (6.9), $\sum_{k=1}^{\infty}\frac{\sigma^{r_k+1}}{\sigma^{r_k+1}+\alpha_k}$ diverges and hence $\sum_{k=1}^{\infty}\frac{1}{\alpha_k} = \infty$. On the other hand, if $\sum_{k=1}^{\infty}\alpha_k^{-1} = \infty$, then we can possibly have three different cases: $\lim_{k\to\infty}\alpha_k = 0$, $\nexists \lim_{k\to\infty}\alpha_k$ or $\lim_{k\to\infty}\alpha_k\in(0,\infty]$. In the first two cases, $\frac{\sigma^{r+1}}{\sigma^{r+1}+\alpha_k} \nrightarrow 0$ for every $\sigma > 0$, and then the corresponding series diverges. In the latter case instead, by Lemma 23, $\sum_{k=1}^{n}\frac{1}{\alpha_k} \sim c_{\sigma,r}\sum_{k=1}^{n}\frac{\sigma^{r+1}}{\sigma^{r+1}+\alpha_k}$ for every $\sigma > 0$. Then, by $\sum_{k=1}^{\infty}\alpha_k^{-1} = \infty$, we deduce that $\sum_{k=1}^{\infty}\frac{\sigma^{r_k+1}}{\sigma^{r_k+1}+\alpha_k}$ diverges for every $\sigma > 0$ and the NSIWT method converges.

(2) We can assume that $0 < \sigma < 1$. For $\sigma = 1$ the result is indeed trivial owing to the equivalence

$$\sum_{k=1}^{\infty}\frac{1}{1+\alpha_k} = \infty \Longleftrightarrow \sum_{k=1}^{\infty}\alpha_k^{-1} = \infty \qquad \text{(see the previous point)}.$$

On the other hand, if $\sigma < 1$ then we have $\sigma^{r_n+1}\to 0$ and $\frac{1}{\sigma^{r_n+1}+\alpha_k}\sim\alpha_k^{-1}$, for $n\to\infty$. Therefore, there exists $N = N(\sigma)$ such that $\frac{1}{\sigma^{r_n+1}+\alpha_k} \ge \frac{1}{2}\alpha_k^{-1}$ for every $n \ge N$. Hence, we have

$$\frac{1}{2}\sigma^{r_n+1}\sum_{k=N}^{n}\alpha_k^{-1} \le \sigma^{r_n+1}\left(\sum_{k=1}^{N-1}\frac{1}{\sigma^{r_n+1}+\alpha_k} + \frac{1}{2}\sum_{k=N}^{n}\alpha_k^{-1}\right) \le \sum_{k=1}^{n}\frac{\sigma^{r_n+1}}{\sigma^{r_n+1}+\alpha_k} \le \sum_{k=1}^{n}\frac{\sigma^{r_k+1}}{\sigma^{r_k+1}+\alpha_k}.$$

Since, by Lemma 22, $\sum_{k=N}^{n}\alpha_k^{-1}\sim\sum_{k=1}^{n}\alpha_k^{-1}$ then, by the preceding inequalities, the hypothesis $\left(\sum_{k=1}^{n}\alpha_k^{-1}\right)^{-1} = o(\sigma^{r_n+1})$ implies that $\sum_{k=1}^{n}\frac{\sigma^{r_k+1}}{\sigma^{r_k+1}+\alpha_k} = \infty$ and the NSIWT method converges. □

Corollary 25 applies immediately to the stationary case, where $\alpha_k = \alpha$ and $r_k = r$ for every $k\in\mathbb{N}$, showing that SIWT converges. On the other hand, from point *(2)* of Corollary 25, given a monotone divergent sequence $r_k \to \infty$ we need a sequence $\alpha_k \to 0$ such that $\alpha_k = o(\sigma^{r_k+1})$ for every $\sigma > 0$ in order to preserve the convergence of NSIWT.

Now, we investigate the convergence rate of NSIWT.

Theorem 26 *Let $\{x^n_{\alpha_n,r_n}\}_{n\in\mathbb{N}}$ be a convergent sequence of the NSIWT method, with $x^\dagger \in \mathcal{X}_\nu$ for some $\nu > 0$, and let $\{\vartheta_n\}_{n\in\mathbb{N}}$ be a divergent sequence of positive real numbers. If*

$$\lim_{n\to\infty} \vartheta_n \sigma^\nu \prod_{k=1}^{n} \left(1 - \frac{\sigma^{r_k+1}}{\sigma^{r_k+1} + \alpha_k}\right) = 0 \qquad \text{for every } \sigma \in \sigma(K) \setminus \{0\}; \tag{6.10a}$$

$$\sup_{\sigma\in\sigma(K)\setminus\{0\}} \vartheta_n \sigma^\nu \prod_{k=1}^{n} \left(1 - \frac{\sigma^{r_k+1}}{\sigma^{r_k+1} + \alpha_k}\right) \le c < \infty \qquad \text{uniformly with respect to } n, \tag{6.10b}$$

then

$$\|x^\dagger - x^n_{\alpha_n,r_n}\| = o(\vartheta_n^{-1}). \tag{6.11}$$

Proof. From equation (6.4), for $x^\dagger \in \mathcal{X}_\nu$, we have

$$\begin{aligned}\lim_{n\to\infty} \vartheta_n \|x^\dagger - x^n_{\alpha_n,r_n}\| &= \lim_{n\to\infty} \left[\int_{\sigma(K^*K)} \left|\vartheta_n \sigma^\nu \prod_{k=1}^{n} \left(1 - \frac{\sigma^{r_k+1}}{\sigma^{r_k+1} + \alpha_k}\right)\right|^2 d\langle E_{\sigma^2}\omega, \omega\rangle\right]^{1/2} \\ &= \left[\int_{\sigma(K^*K)} \left|\lim_{n\to\infty} \vartheta_n \sigma^\nu \prod_{k=1}^{n} \left(1 - \frac{\sigma^{r_k+1}}{\sigma^{r_k+1} + \alpha_k}\right)\right|^2 d\langle E_{\sigma^2}\omega, \omega\rangle\right]^{1/2},\end{aligned}$$

by (6.10b) and the Dominated Convergence Theorem. Now, from hypothesis (6.10a), the thesis follows. □

Corollary 27 *We define*

$$\beta_n = \sum_{k=1}^{n} \alpha_k^{-1}, \qquad \tilde{\beta}_n = \sum_{k=1}^{n} \frac{1}{1+\alpha_k}.$$

Let $\{r_k\}_{k\in\mathbb{N}}$ be a sequence of positive real numbers, $r_k \ge 0$, and let $x^\dagger \in \mathcal{X}_\nu$ for some $\nu > 0$. If

(i.1) $\sup_{k\in\mathbb{N}}\{r_k\} = r \in [0, \infty)$,

(i.2) $\lim_{n\to\infty} \beta_n = \infty$,

then

$$\|x^\dagger - x^n_{\alpha_n,r_n}\| = \begin{cases} o(\beta_n^{-\frac{\nu}{r+1}}) & \text{if } \lim_{n\to\infty} \alpha_n = \alpha \in (0, \infty] \quad (6.12a) \\ O(\beta_n^{-\frac{\nu}{r+1}}) & \text{if } \lim_{n\to\infty} \alpha_n = 0 \text{ and } \alpha_n^{-1} \le c\beta_{n-1},\ c > 0 \quad (6.12b) \\ o(\tilde{\beta}_n^{-\frac{\nu}{r+1}}) & \text{otherwise.} \quad (6.12c) \end{cases}$$

On the contrary, if

(ii.1) $r_k \to \infty$ *monotonically,*

(ii.2) $\beta_n^{-1} = o(\sigma^{r_n+1})$ *for every* $\sigma \in \sigma(K) \setminus \{0\}$*,*

then

$$\|x^\dagger - x^n_{\alpha_n,r_n}\| = o(\beta_n^{-\frac{\nu}{r_n+1}}). \tag{6.13}$$

Proof. First, note that from (i.1), (i.2) and Corollary 25 it follows that the NSIWT method is convergent. Now, since $1 - x \le \mathrm{e}^{-x} \le c_{\nu,r} x^{-\nu/r+1}$, and using (i.2), we have

$$\begin{aligned}
\sigma^\nu \prod_{k=1}^n \left(1 - \frac{\sigma^{r_k+1}}{\sigma^{r_k+1} + \alpha_k}\right) &\le \sigma^\nu \mathrm{e}^{-\sum_{k=1}^n \frac{\sigma^{r_k+1}}{\sigma^{r_k+1}+\alpha_k}} \\
&\le \sigma^\nu \mathrm{e}^{-\sigma^{r+1}\sum_{k=1}^n \frac{1}{\sigma^{r+1}+\alpha_k}} \\
&\le c_{\nu,r}\sigma^\nu \left(\frac{1}{\sigma^{r+1}\sum_{k=1}^n \frac{1}{\sigma^{r+1}+\alpha_k}}\right)^{\frac{\nu}{r+1}} \\
&\le c_{\nu,r} \left(\sum_{k=1}^n \frac{1}{1+\alpha_k}\right)^{-\frac{\nu}{r+1}}.
\end{aligned}$$

Therefore, conditions (6.10a) and (6.10b) of Theorem 26 are satisfied with $\vartheta_n = \left(\sum_{k=1}^n \frac{1}{1+\alpha_k}\right)^{\frac{\nu}{r+1}}$. If $\lim_{k\to\infty} \alpha_k = \alpha \in (0, \infty]$, then $\beta_n \sim c\sum_{k=1}^n \frac{1}{1+\alpha_k}$ for $n \to \infty$ by Lemma 23. Equations (6.12a) and (6.12c) follow. Eventually, observing that $1 - \frac{\sigma^{r_k+1}}{\sigma^{r_k+1}+\alpha_k} \le 1 - \frac{\sigma^{r+1}}{\sigma^{r+1}+\alpha_k}$, equation (6.12b) follows instead by a straightforward application of [Lemma 1,2,3 and Theorem 1] [8].

To prove equation (6.13) the strategy is the same. We have $e^{-x} \le x^{-\nu/(r_n+1)}$ definitely, $1/(\sigma^{r_n+1} + \alpha_k) \sim \alpha_k^{-1}$ for $n \to \infty$, and hypothesis (ii.2) implies that $\beta_n^{-1/(r_n+1)} \to 0$ converges to zero. □

When $r = 1$ (classical iterated Tikhonov), equation (6.12b) is the result in [8, Theorem 1]. On the other hand, if $\lim_{n\to\infty} \alpha_n = \alpha \in (0, \infty]$, then the convergence rate is improved by the small "o".

Remark 28 *As we stated in* (6.12b)*, when* $\lim_{n\to\infty} \alpha_n = 0$*, to obtain a convergence rate of order* $O(\beta_n^{-\nu/(r+1)})$ *the sequence* $\{\alpha_n\}$ *has to satisfy the condition* $\alpha_n^{-1} \le c\beta_{n-1}$ *for a positive real number* $c > 0$*. Then,* $\sum_{k=1}^n \alpha_k^{-1} = \beta_n = O(q^n)$*, where* $q = (1 + c) > 1$*. To overcome this bound, in virtue of (ii.1), (ii.2) of Corollary 27, choosing sequences* $\{\hat{r}_n\}$ *and* $\{\hat{\alpha}_n\}$ *such that* $\hat{r}_n$ *diverges monotonically and* $\left(\sum_{k=1}^n \hat{\alpha}_k^{-1}\right)^{-1} = o(\sigma^{\hat{r}_n+1})$ *for every* $0 < \sigma \le 1$*, we are able to obtain a faster convergence rate, in a sense that has still to be defined. In the following Proposition 29 we will give the proof for a specific case.*

Following the same approach in [1, (2.3), (2.4) pag. 26], we say that the sequence $\{\hat{x}_n\}$ converges uniformly faster than the sequence $\{x_n\}$ if

$$x^\dagger - \hat{x}_n = R_n(x^\dagger - x_n), \tag{6.14}$$

where $\{R_n\}$ is a sequence of operators such that $\|R_n\| \to 0$ as $n \to \infty$. We say instead that $\{\hat{x}_n\}$ converges non-uniformly faster than $\{x_n\}$ if (6.14) holds and

$$\inf_{n\in\mathbb{N}} \|R_n\| > 0, \qquad \lim_{n\to\infty} \|R_n x\| = 0 \text{ for every } x \in \mathcal{X}.$$

We are ready to state the following comparison result.

Proposition 29 *Let $\{x^n_{\alpha_n}\}$ be the sequence generated by the nonstationary iterated Tikhonov with $\alpha_n = \alpha_0 q^n$, where $\alpha_0 \in (0,\infty), q \in (0,1)$, and let $\{x^n_{\hat{\alpha}_n,\hat{r}_n}\}$ be the sequence generated by NSIWT, where $\hat{\alpha}_n = 1/n!$ and $\hat{r}_n = n$, both applied to the same compact operator $K : \mathcal{X} \to \mathcal{Y}$. Then, $\{x^n_{\hat{\alpha}_n,\hat{r}_n}\}$ converges, non uniformly, faster than $\{x^n_{\alpha_n}\}$.*

Proof. Observe that the sequence $\{x^n_{\alpha_n}\}$ corresponds to a NSIWT method $\{x^n_{\alpha_n,r_n}\}$ with $r_n = 1$ for every n. Moreover, both the sequences $\{x^n_{\alpha_n}\}$ and $\{x^n_{\hat{\alpha}_n,\hat{r}_n}\}$ converge, indeed they satisfy conditions (1) and (2) of Corollary 25, respectively. Assuming that $x_0 = 0$ and applying the same strategy used in Theorem 24, without any effort it is possible to show that

$$x^\dagger - x^n_{\hat{\alpha}_n,\hat{r}_n} = \prod_{k=1}^n \hat{\alpha}_k \left((K^*K)^{\frac{\hat{r}_k+1}{2}} + \hat{\alpha}_k I\right)^{-1} x^\dagger,$$

$$x^\dagger = \prod_{k=1}^n \alpha_k^{-1} (K^*K + \alpha_k I)(x^\dagger - x^n_{\alpha_n}).$$

Therefore we find

$$x^\dagger - x^n_{\hat{\alpha}_n,\hat{r}_n} = \left[\prod_{k=1}^n \hat{\alpha}_k \alpha_k^{-1} \left((K^*K)^{\frac{\hat{r}_k+1}{2}} + \hat{\alpha}_k I\right)^{-1} (K^*K + \alpha_k I)\right](x^\dagger - x^n_{\alpha_n}) = R_n(x^\dagger - x^n_{\alpha_n}).$$

Since $0 \in \sigma(K^*K)$, we infer $\|R_n\| > 1$ for every n, and hence $\inf_{n\in\mathbb{N}} \|R_n\| \geq 1$. If we prove that

$$\lim_{n\to\infty} \|R_n x\| = 0,$$

for every $x \in \mathcal{X}$, then the thesis follows. Since

$$\lim_{n\to\infty} \|R_n x\| = 0 \Longleftrightarrow \lim_{n\to\infty} \prod_{k=1}^n \frac{\hat{\alpha}_k(\sigma^2 + \alpha_k)}{\alpha_k(\sigma^{\hat{r}_k+1} + \hat{\alpha}_k)} = 0 \Longleftrightarrow \sum_{k=1}^\infty \frac{\alpha_k \sigma^{\hat{r}_k+1} - \hat{\alpha}_k \sigma^2}{\alpha_k \sigma^{\hat{r}_k+1} + \alpha_k \hat{\alpha}_k} = \infty \;\; \forall \sigma > 0,$$

if we substitute the values $\alpha_n = \alpha_0 q^n$, then $\hat{\alpha}_n = 1/n!$ and $\hat{r}_n = n$, we obtain

$$\sum_{k=1}^\infty \frac{\alpha_k \sigma^{\hat{r}_k+1} - \hat{\alpha}_k \sigma^2}{\alpha_k \sigma^{\hat{r}_k+1} + \alpha_k \hat{\alpha}_k} = \sum_{k=1}^\infty \frac{1 - \frac{\sigma}{\alpha_0 n!(q\sigma)^n}}{1 + \frac{1/n!}{\sigma^{n+1}}},$$

and the right hand side of the above equality diverges: indeed

$$\frac{1 - \frac{\sigma}{\alpha_0 n!(q\sigma)^n}}{1 + \frac{1/n!}{\sigma^{n+1}}} \longrightarrow 1 \text{ for every fixed } q, \sigma \in (0,1) \text{ and } \alpha_0 \in (0,\infty).$$

□

6.2 Analysis of convergence for perturbed data

Let now consider $y^\delta = y + \delta\eta$, with $y \in R(K)$ and $\|\eta\| = 1$, i.e., $\|y^\delta - y\| = \delta$. We are concerned about the convergence of the NSIWT method when the initial datum y is perturbed. Hereafter we will use the notation $x^{n,\delta}_{\alpha_n,r_n}$ for the solution of NSIWT (6.2) with initial datum y^δ.

The following result can be proved similarly to Theorem 1.7 in [1].

Theorem 30 *Under the assumptions of Corollary 25, if $\{\delta_n\}$ is a sequence convergent to 0 with $\delta_n \geq 0$ and such that*

$$\lim_{n\to\infty} \delta_n \cdot \sum_{k=1}^{n} \alpha_k^{-1} = 0, \tag{6.15}$$

then, $\lim_{n\to\infty} \|x^\dagger - x^{n,\delta_n}_{\alpha_n,r_n}\| = 0$.

Proof. From the definition of the method (6.1), for every given j, n, we find that

$$\begin{aligned}
x^{j,\delta_n}_{\alpha_j,r_j} &= \left[(K^*K)^{\frac{r_j+1}{2}} + \alpha_j I\right]^{-1} \left((K^*K)^{\frac{r_j-1}{2}} K^* y^{\delta_n} + \alpha_j x^{j-1,\delta_n}_{\alpha_{j-1},r_{j-1}}\right) \\
&= \left\{I - \alpha_j \left[(K^*K)^{\frac{r_j+1}{2}} + \alpha_j I\right]^{-1}\right\} x^\dagger + \alpha_j \left[(K^*K)^{\frac{r_j+1}{2}} + \alpha_j I\right]^{-1} x^{j-1,\delta_n}_{\alpha_{j-1},r_{j-1}} \\
&\quad + \left[(K^*K)^{\frac{r_j+1}{2}} + \alpha_j I\right]^{-1} (K^*K)^{\frac{r_j-1}{2}} K^*(y^{\delta_n} - y),
\end{aligned}$$

namely,

$$\begin{aligned}
x^\dagger - x^{j,\delta_n}_{\alpha_j,r_j} &= \alpha_j \left[(K^*K)^{\frac{r_j+1}{2}} + \alpha_j I\right]^{-1} (x^\dagger - x^{j-1,\delta_n}_{\alpha_{j-1},r_{j-1}}) \\
&\quad - \left[(K^*K)^{\frac{r_j+1}{2}} + \alpha_j I\right]^{-1} (K^*K)^{\frac{r_j-1}{2}} K^*(y^{\delta_n} - y).
\end{aligned}$$

Hence, by induction, for every fixed n we have

$$\begin{aligned}
x^\dagger - x^{n,\delta_n}_{\alpha_n,r_n} &= \prod_{k=1}^{n} \alpha_k \left[(K^*K)^{\frac{r_k+1}{2}} + \alpha_k I\right]^{-1} x^\dagger \\
&\quad - \sum_{k=1}^{n} \alpha_k^{-1} \prod_{i=k}^{n} \alpha_i \left[(K^*K)^{\frac{r_i+1}{2}} + \alpha_i I\right]^{-1} (K^*K)^{\frac{r_k-1}{2}} K^*(y^{\delta_n} - y).
\end{aligned}$$

If we set $g_{k,n}(K^*K) = \prod_{i=k}^{n} \alpha_i \left[(K^*K)^{\frac{r_i+1}{2}} + \alpha_i I\right]^{-1} (K^*K)^{\frac{r_k-1}{2}}$, then we have

$$\begin{aligned}
\|g_{k,n}(K^*K)K^*y\|^2 &= \langle g_{k,n}(K^*K)K^*y, g_{k,n}(K^*K)K^*y\rangle \\
&= \langle g_{k,n}(KK^*)KK^*y, g_{k,n}(KK^*)y\rangle \\
&= \langle g_{k,n}(KK^*)(KK^*)^{1/2}y, g_{k,n}(K^*K)(KK^*)^{1/2}y\rangle \\
&= \|g_{k,n}(KK^*)(KK^*)^{1/2}y\|^2,
\end{aligned}$$

where we used the fact that $g_{k,n}(K^*K)K^* = K^*g_{k,n}(KK^*)$ and that for every bounded Borel function f and h, the product $f(A)h(B)$ commutes if the self-adjoint operators A and B commute [14, see 12.24]. Therefore,

$$\left\| \prod_{j=k}^{n} \alpha_j \left[(K^*K)^{\frac{r_j+1}{2}} + \alpha_j I \right]^{-1} (K^*K)^{\frac{r_k-1}{2}} K^* \right\| = \left\| \prod_{j=k}^{n} \alpha_j \left[(KK^*)^{\frac{r_j+1}{2}} + \alpha_j I \right]^{-1} (KK^*)^{\frac{r_k}{2}} \right\|$$
$$= \max_{\sigma\in[0,1]} \left| \sigma^{r_k} \prod_{j=k}^{n} \frac{\alpha_j}{\sigma^{r_j+1} + \alpha_j} \right| \leq 1.$$

It follows that

$$\|x^\dagger - x^{n,\delta_n}_{\alpha_n,r_n}\| \leq \| \prod_{k=1}^{n} \alpha_k \left[(K^*K)^{\frac{r_k+1}{2}} + \alpha_k I \right]^{-1} x^\dagger \| + \sum_{k=1}^{n} \alpha_k^{-1} \|y^{\delta_n} - y\|$$
$$= \|x^\dagger - x^n_{\alpha_n,r_n}\| + \delta_n \sum_{k=1}^{n} \alpha_k^{-1},$$

and by Corollary 27 and (6.15), $\|x^\dagger - x^{n,\delta_n}_{\alpha_n,r_n}\| \to 0$ for $n \to \infty$. □

7 Nonstationary iterated fractional Tikhonov

Definition 31 (Nonstationary iterated fractional Tikhonov) *Let $\{\alpha_n\}_{n\in\mathbb{N}}$ and $\{\gamma_n\}_{n\in\mathbb{N}}$ be sequences of real numbers such that $\alpha_n > 0$ and $\gamma_n \geq 1/2$ for every n. We define the* nonstationary iterated fractional Tikhonov method (NSIFT) *as*

$$\begin{cases} x^0_{\alpha_0,\gamma_0} := 0; \\ (K^*K + \alpha_n I)^{\gamma_n} x^n_{\alpha_n,\gamma_n} := (K^*K)^{\gamma_n - 1} K^* y + \left[(K^*K + \alpha_n I)^{\gamma_n} - (K^*K)^{\gamma_n} \right] x^{n-1}_{\alpha_{n-1},\gamma_{n-1}}. \end{cases} \tag{7.1}$$

We denote by $x^{n,\delta}_{\alpha_n,\gamma_n}$ the n-th iteration of NSIFT if $y = y^\delta$.

Theorem 32 *The* NSIFT *method* (7.1) *converges to $x^\dagger \in \mathcal{X}$ as $n \to \infty$ if and only if $\sum_n \left(\frac{\sigma^2}{\sigma^2+\alpha_n} \right)^{\gamma_n}$ diverges for every $\sigma > 0$.*

Proof. The proof follows the same steps as in Theorem 24. Therefore we will omit details. What follows is that

$$x^\dagger - x^n_{\alpha_n,\gamma_n} = \prod_{k=1}^{n} (K^*K + \alpha_k I)^{-\gamma_k} \left[(K^*K + \alpha_k I)^{\gamma_k} - (K^*K)^{\gamma_k} \right] x^\dagger,$$

and hence

$$\|x^\dagger - x^n_{\alpha_n,\gamma_n}\|^2 = \int_{\sigma(K^*K)} \left| \prod_{k=1}^n \frac{(\sigma^2+\alpha_k)^{\gamma_k} - \sigma^{2\gamma_k}}{(\sigma^2+\alpha_k)^{\gamma_k}} \right|^2 d\langle E_{\sigma^2} x^\dagger, x^\dagger \rangle.$$

Then, the method converges if and only if

$$\lim_{n\to\infty} \prod_{k=1}^n \left[1 - \left(\frac{\sigma^2}{\sigma^2+\alpha_k} \right)^{\gamma_k} \right] = 0$$

for every $\sigma > 0$. The thesis follows by Lemma 21. □

Corollary 33

(1) Let $\lim_{k\to\infty} \gamma_k = \gamma \in [1/2, \infty)$. Then the NSIFT method converges if and only if

$$\sum_{k=1}^n \alpha_k^{-\gamma} = \infty.$$

More in general, if $\sup_{k\in\mathbb{N}}\{\gamma_k\} = s \in [1/2, \infty)$ and $\sum_{k=1}^\infty \alpha_k^{-s} = \infty$, then the NSIFT method converges.

(2) Let $\lim_{k\to\infty} \gamma_k = \infty$. If $\lim_{k\to\infty} \alpha_k = 0$ and $\lim_{k\to\infty} \alpha_k\gamma_k = l \in [0, \infty)$, then the NSIFT method converges.

Proof. (1) It is immediate noticing that

$$\sum_{k=1}^n \left(\frac{\sigma^2}{\sigma^2+\alpha_k} \right)^{\gamma_k} \sim c \sum_{k=1}^n \left(\frac{\sigma^2}{\sigma^2+\alpha_k} \right)^{\gamma}$$
$$\sum_{k=1}^n \left(\frac{\sigma^2}{\sigma^2+\alpha_k} \right)^{\gamma_k} \geq \sum_{k=1}^n \left(\frac{\sigma^2}{\sigma^2+\alpha_k} \right)^{s}.$$

(2) We observe that

$$\left(\frac{\sigma^2}{\sigma^2+\alpha_k} \right)^{\gamma_k} = \left(1 - \frac{\alpha_k}{\sigma^2+\alpha_k} \right)^{\gamma_k} \sim \mathrm{e}^{-\frac{\alpha_k\gamma_k}{\sigma^2+\alpha_k}} \to \mathrm{e}^{-l/\sigma^2} \neq 0$$

for $k \to \infty$. Then $\sum_{k=1}^n \left(\frac{\sigma^2}{\sigma^2+\alpha_k} \right)^{\gamma_k}$ diverges for every $\sigma > 0$ and the NSIFT method converges. □

Theorem 34 *Let $\{x^n_{\alpha_n,\gamma_n}\}_{n\in\mathbb{N}}$ be a convergent sequence of the NSIFT method, with $x^\dagger \in \mathcal{X}_\nu$ for some $\nu > 0$, and let $\{\vartheta_n\}_{n\in\mathbb{N}}$ be a divergent sequence of positive real numbers. If*

$$\lim_{n\to\infty} \vartheta_n \sigma^\nu \prod_{k=1}^n \left(1 - \frac{\sigma^{2\gamma_k}}{(\sigma^2+\alpha_k)^{\gamma_k}} \right) = 0 \qquad \textit{for every } \sigma \in \sigma(K) \setminus \{0\}; \tag{7.2a}$$

$$\sup_{\sigma\in\sigma(K)\setminus\{0\}} \vartheta_n \sigma^\nu \prod_{k=1}^{n}\left(1-\frac{\sigma^{2\gamma_k}}{(\sigma^2+\alpha_k)^{\gamma_k}}\right) \le c < \infty \qquad \textit{uniformly with respect to } n, \tag{7.2b}$$

then

$$\|x^\dagger - x^n_{\alpha_n,\gamma_n}\| = o(\vartheta_n^{-1}). \tag{7.3}$$

Proof. As seen in Theorem 26, the thesis follows easily from the Dominated Convergence Theorem. □

Corollary 35 *Let* $\{\gamma_k\}_{k\in\mathbb{N}}$ *be a sequence of positive real numbers,* $\gamma_k \ge 1/2$*, and let* $x^\dagger \in \mathcal{X}_\nu$ *for some* $\nu > 0$*. If*

(i.1) $\sup_{k\in\mathbb{N}}\{\gamma_k\} = s \in [1/2,\infty)$,

(i.2) $\lim_{n\to\infty}\beta_n = \infty$,

then

$$\|x^\dagger - x^n_{\alpha_n,\gamma_n}\| = o(\beta_n^{-\frac{\nu}{2s}}) \qquad \textit{if } \exists \lim_{k\to\infty}\alpha_k = \alpha \in (0,\infty], \tag{7.4}$$

$$\|x^\dagger - x^n_{\alpha_n,\gamma_n}\| = o(\tilde\beta_n^{-\frac{\nu}{2s}}) \qquad \textit{otherwise}, \tag{7.5}$$

where we defined

$$\beta_n = \sum_{k=1}^{n}\alpha_k^{-s}, \qquad \tilde\beta_n = \sum_{k=1}^{n}\frac{1}{1+\alpha_k^s}.$$

On the contrary, if

(ii.1) $\lim_{k\to\infty}\gamma_k = \infty$,

(ii.2) $\lim_{k\to\infty}\alpha_k = 0$ *and* $\lim_{k\to\infty}\alpha_k\gamma_k = 0$,

then

$$\|x^\dagger - x^n_{\alpha_n,\gamma_n}\| = o(n^{-1}). \tag{7.6}$$

Proof. See Corollary 27. In particular, for the second statement we use the fact that

$$\mathrm{e}^{-\sum_{k=1}^{n}\left(\frac{\sigma^2}{\sigma^2+\alpha_k}\right)^{\gamma_k}} = o(n^{-1}).$$

□

Theorem 36 *Under the assumptions of Corollary 33, if* $\{\delta_n\}$ *is a sequence convergent to* 0 *with* $\delta_n \ge 0$ *and such that*

$$\lim_{n\to\infty}\delta_n\cdot\sum_{k=1}^{n}\alpha_k^{-\gamma_k} = 0, \tag{7.7}$$

then, $\lim_{n\to\infty}\|x^\dagger - x^{n,\delta_n}_{\alpha_n,\gamma_n}\| = 0$.

Proof. Here is a sketch of the proof, since it follows step by step from the proof of Theorem 30. If we set

$$\psi_k(K^*K) := [(K^*K + \alpha_k I)^{\gamma_k} - (K^*K)^{\gamma_k}]$$
$$\phi_k(K^*K) := \psi_k(K^*K)\,[K^*K + \alpha_k I]^{-\gamma_k}\,,$$

then from (7.1) it is possible to show that

$$x^\dagger - x^{n,\delta_n}_{\alpha_n,\gamma_n} = \prod_{k=1}^{n} \phi_k(K^*K)x^\dagger - \sum_{k=1}^{n} \psi_k(K^*K)^{-1} \prod_{i=k}^{n} \phi_i(K^*K)\,(K^*K)^{\gamma_k - 1}\, K^*(y^{\delta_n} - y),$$

for every integer n and for every perturbed data $y^{\delta_n} = y + \delta_n \eta$. Owing to the equality

$$\left\| \prod_{i=k}^{n} \phi_i(K^*K)\,(K^*K)^{\gamma_k-1}\,K^* \right\| = \left\| \prod_{i=k}^{n} \phi_i(KK^*)\,(KK^*)^{\gamma_k-1}\,(KK^*)^{1/2} \right\|,$$

we deduce

$$\begin{aligned} \|x^\dagger - x^{n,\delta_n}_{\alpha_n,\gamma_n}\| &\leq \|x^\dagger - x^{n}_{\alpha_n,\gamma_n}\| + \delta_n \sum_{k=1}^{n} \left\|\psi_k(K^*K)^{-1}\right\| \\ &= \|x^\dagger - x^{n}_{\alpha_n,\gamma_n}\| + \delta_n \sum_{k=1}^{n} \alpha_k^{-\gamma_k}. \end{aligned}$$

□

8 Numerical results

We now give few selected examples with a special focus on the nonstationary iterations proposed in this paper. For a larger comparison between fractional and classical Tikhonov refer to [6,7,9]. To produce our results we used Matlab 8.1.0.604 using a laptop pc with processor Intel iCore i5-3337U with 6 GB of RAM running Windows 8.1.

We add to the noise-free right-hand side vector y, the "noise-vector" e that has in all examples normally distributed pseudorandom entries with mean zero, and is normalized to correspond to a chosen noise-level

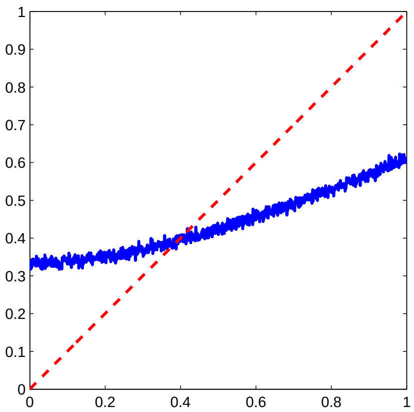
$$\xi = \frac{\|e\|}{\|y\|}.$$

As a stopping criterion for the methods we used the Discrepancy Principle [12], that terminates the iterative method at the iteration

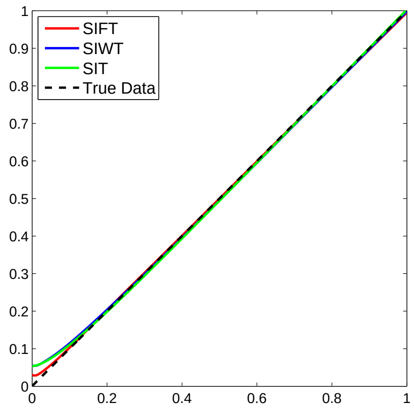
$$\hat{k} = \min_k \{k : \ \|y^\delta - Kx_k\| \leq \tau\delta\},$$

where $\tau = 1.01$. This criterion stops the iterations when the norm of the residual reaches the norm of the noise so that the latter is not reconstructed.

To compare the restorations with the different methods, we consider both the visual representation and the relative restoration error that is $\|\hat{x} - x^\dagger\|/\|x^\dagger\|$ for the computed approximation $\hat{x}$.



(a)



(b)

Figure 1: Example 1 – "Foxgood" test case: (a) the true solution (dashed curve) and the observed data (solid curve), (b) approximated solutions by SIFT with $\gamma = 0.8$ and $\alpha = 10^{-3}$, SIWT with $r = 0.6$ and $\alpha = 10^{-2}$, and SIWT with $r = 1$ and $\alpha = 10^{-3}$.

8.1 Example 1

This test case is the so-called *Foxgood* in the toolbox Regularization tool by P. Hansen [2] using 1024 points. We have added a noise vector with $\xi = 0.02$ to the observed signal. In Figure 1(a) the true signal and the measured data can be seen.

In Table 1 we show the relative errors with different choices of α, r and γ. In brackets we report the iteration at which the discrepancy principle stopped the method. Note that SIFT with $\gamma = 1$ and SIWT with $r = 1$ are exactly the classical Tikhonov method and hence produce the same result. Figure 1(b) shows the reconstruction for SIFT with $\gamma = 0.8$ and $\alpha = 10^{-3}$, SIWT with $r = 0.6$ and $\alpha = 10^{-2}$, and SIWT with $r = 1$ (classical Iterated Tikhonov) with $\alpha = 10^{-3}$.

From these results, using both fractional and weighted iterated Tikhonov, we can see that we can obtain better restorations than with the classical version. However, in order to obtain such results, one has to evaluate α very carefully. Indeed α does not only affects the convergence speed, but also the quality of the restoration: a small perturbation in α can lead to quite different restoration errors. The nonstationary version of the methods can help also to avoid such a careful and often difficult estimation.

For the nonstationary iterations we assume the regularization parameter α_n at each iteration be given according to the geometric sequence

$$\alpha_n = \alpha_0 q^n, \qquad q \in (0,1), \qquad n = 1, 2, \ldots. \tag{8.1}$$

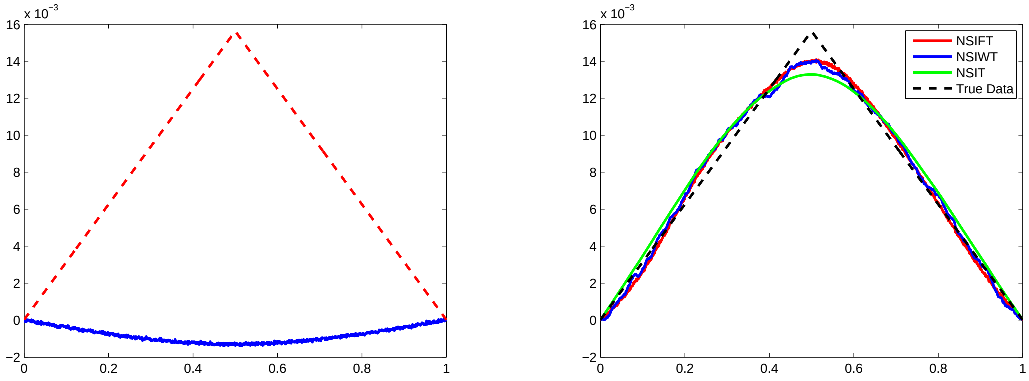
Setting $r_n = 0.6$ and $\gamma_n = 0.8$, Table 2 shows that NSIFT and NSIWT provide a relative error lower than the classical nonstationary iterated Tikhonov (NSIT). Finally, since NSIFT and NSIWT allow a nonstationary choice also for r_n and γ_n, in Table 2 we report the results

α	Method	r/γ				
		0.4	0.6	0.8	1	1.2
5×10^{-2}	SIFT	337.09(7)	0.02498(13)	0.03481(19)	0.03752(29)	0.03838(43)
	SIWT	0.02589(9)	0.03202(13)	0.03609(19)	0.03752(29)	0.03932(43)
10^{-2}	SIFT	320.85(3)	0.02048(5)	0.02633(7)	0.03731(7)	0.03783(9)
	SIWT	0.01697(3)	0.01818(5)	0.03361(5)	0.03731(7)	0.03672(11)
5×10^{-3}	SIFT	423.37(3)	0.02216(3)	0.02190(5)	0.03102(5)	0.03723(5)
	SIWT	0.02421(3)	0.01573(3)	0.03186(3)	0.03103(5)	0.03347(7)
10^{-3}	SIFT	402.97(1)	0.02299(1)	**0.00698(3)**	0.01756(3)	0.02443(3)
	SIWT	0.06403(1)	0.02210(1)	0.02528(1)	0.01756(3)	0.02736(3)
5×10^{-4}	SIFT	531.72(1)	0.02119(1)	0.01729(1)	0.02507(1)	0.03119(1)
	SIWT	0.10518(1)	0.04506(1)	0.01482(1)	0.02507(1)	0.02086(3)
10^{-4}	SIFT	1012.2(1)	0.07246(1)	0.04229(1)	0.02704(1)	0.01675(1)
	SIWT	0.25927(1)	0.13000(1)	0.07213(1)	0.02704(1)	0.01154(1)

Table 1: Example 1: relative errors for SIWT and SIFT for different choices of α, r, and γ.

α_0	Method	q		
		0.7	0.8	0.9
10^{-1}	NSIFT ($\gamma_n = 0.8$)	0.024453(9)	0.030868(11)	0.028849(17)
	NSIWT ($r_n = 0.6$)	0.025223(7)	0.027628(9)	0.028534(13)
	NSIT	0.035162(9)	0.031627(13)	0.036472(19)
	NSIFT (γ_n in (8.2))	0.032489(9)	0.027974(13)	0.037199(17)
	NSIWT (r_n in (8.2))	0.031493(9)	0.027436(13)	0.036059(17)
10^{-2}	NSIFT ($\gamma_n = 0.8$)	**0.014781(5)**	0.021687(5)	0.028709(5)
	NSIWT ($r_n = 0.6$)	**0.014503(3)**	0.021501(3)	0.028396(3)
	NSIT	0.024838(5)	0.030866(5)	0.028835(7)
	NSIFT (γ_n in (8.2))	0.023848(5)	0.030002(5)	0.027636(7)
	NSIWT (r_n in (8.2))	0.023482(5)	0.029638(5)	0.027366(7)

Table 2: Example 1: relative errors for NSIWT and NSIFT with the nonstationary α_n in (8.1) and different choices of r_n and γ_n (NSIT is $r_n = \gamma_n = 1$).



(a) (b)

Figure 2: Example 2 – “deriv2” test case: (a) the true solution (dashed curve) and the observed data (solid curve), (b) approximated solutions.

for the following nonincreasing sequences

$$r_n = \gamma_n = \begin{cases} 1 - \frac{n-1}{100} & n < 50, \\ \frac{1}{2} & \text{otherwise.} \end{cases} \tag{8.2}$$

Again both NSIWT and NSIFT are able to get better results than NSIT. Even tough the errors are not as good as those for the best choices $r_n = 0.6$ and $\gamma_n = 0.8$, the choice (8.2) stresses the robustness of our nonstationary iterations.

8.2 Example 2

We consider the test problem *deriv2(·,3)* in the toolbox REGULARIZATION TOOL by P. Hansen [2] using 1024 points. For the noise vector it holds $\xi = 0.05$. In Figure 2(a) we can see the measured data and the true signal. We compare NSIWT and NSIFT with the NSIT.

Firstly, α_n is defined by the classical choice in (8.1). Table 3 shows the results for different choices of r_n and γ_n. Note that NSIWT and NSIFT usually outperform NSIT. Nevertheless, our nonstationary iterations allow also unbounded sequences of r_n and γ_n. Therefore, according to Proposition 29, we set

$$\alpha_n = \frac{1}{n!}, \qquad r_n = \frac{n}{10}, \qquad \gamma_n = \frac{n}{2}. \tag{8.3}$$

Table 4 shows that the relative restoration error obtained with the unbounded sequences r_n and γ_n in (8.3) is lower than the best one (according to Table 3), obtained by NSIT by employing the geometric sequence (8.1) for α_n. The computed approximations are also compared in Figure 2(b), where we note a better restoration of the corner for NSIWT and NSIFT.

8.3 Example 3

We consider the test problem *blur(·,·,·)* in the toolbox REGULARIZATION TOOL by P. Hansen [2]. This is a two dimensional deblurring problem, the true solution is a 40×40 image, the blurring

α_0	Method	q		
		0.7	0.8	0.9
10^{-1}	NSIFT ($\gamma_n = 0.8$)	0.08981(11)	0.09394(13)	0.09445(19)
	NSIWT ($r_n = 0.6$)	0.08051(13)	0.09181(17)	0.09401(29)
	NSIT	0.08502(15)	0.09175(21)	0.09466(37)
	NSIFT (γ_n in (8.2))	0.09428(13)	0.09089(19)	0.09327(29)
	NSIWT (r_n in (8.2))	0.09073(13)	0.08648(19)	0.09199(29)
10^{-2}	NSIFT ($\gamma_n = 0.8$)	0.09114(5)	0.08953(7)	0.08998(9)
	NSIWT ($r_n = 0.6$)	**0.07807(7)**	0.09411(7)	0.09183(11)
	NSIT	0.08183(9)	0.09174(11)	0.09379(17)
	NSIFT (γ_n in (8.2))	**0.07839(9)**	0.08721(11)	0.09246(15)
	NSIWT (r_n in (8.2))	0.09399(7)	0.08389(11)	0.08990(15)

Table 3: Example 2: relative errors for NSIWT and NSIFT with the nonstationary α_n in (8.1) and different choices of r_n and γ_n (NSIT is $r_n = \gamma_n = 1$).

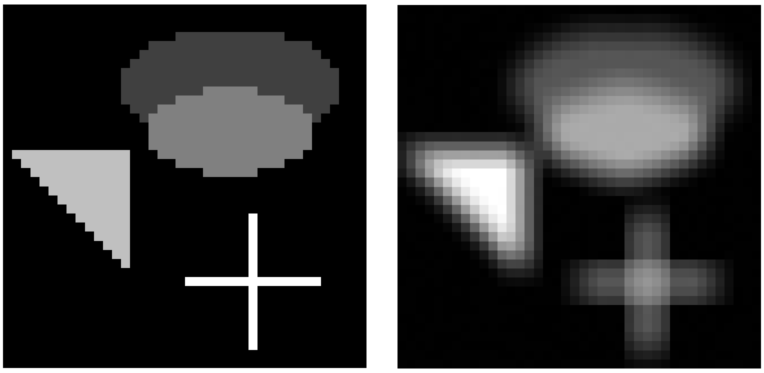
	NSIFT	NSIWT	NSIT
Error	0.054831(9)	0.059211(7)	0.081835(9)

Table 4: Example 2: relative restorations errors for NSIFT and NSIWT with parameters in (8.3) and NSIT with $\alpha_n = 0.01 \cdot 0.7^n$.

operator is a symmetric BTTB (block Toeplitz with Toeplitz block) with bandwidth 6. This blur is created by a truncated Gaussian point spread function with variance 2. For the noise vector it holds $\nu = 0.005$. Figure 3(a) shows the true image while the observed image is in Figure 3(b).

Firstly, α_n is defined by the classical choice in (8.1). Table 5 provides the results for a good stationary choice of r_n and γ_n. Note that NSIWT and NSIFT usually outperform NSIT. Table 6 shows that the relative restoration error obtained with the unbounded sequences r_n and γ_n in (8.3) is lower than the best one (according to Table 5), obtained by the stationary choice of r_n and γ_n. We note that using r_n and γ_n in (8.3), we are able to get better reconstruction than NSIT with the geometric sequence for α_n, in particular of the edges, see Figure 4.

Finally, note that for the NSIT a nondecreasing sequence of α_n could be considered instead of the geometric sequence (8.1), see [11]. Nevertheless, this strategy requires a proper choice of α_0 and this is out of the scope of this paper, but it could be investigated in the future in connection with our fractional and weighted variants. A further development of our iterative schemes is in the direction of the nonstationary preconditioning strategy in [5], which is inspired by an approximated solution of the NSIT and hence could be investigated also in a fractional framework.



(a) (b)

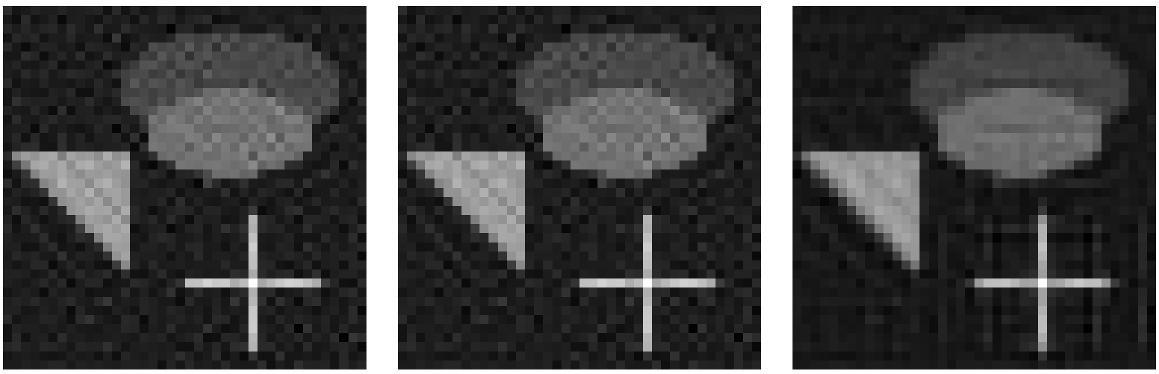
Figure 3: Example 3 – "blur" test case: (a) the true image, (b) the measured data.

α_0	Method	q		
		0.7	0.8	0.9
	NSIFT ($\gamma_n = 0.5$)	0.19970(9)	0.19526(13)	0.19847(17)
10^{-1}	NSIWT ($r_n = 0.2$)	0.18936(7)	**0.18920(9)**	0.19732(11)
	NSIT	0.19816(15)	0.21786(20)	0.28703(20)
	NSIFT ($\gamma_n = 0.5$)	0.19398(5)	0.19962(5)	0.19595(7)
10^{-2}	NSIWT ($r_n = 0.2$)	0.20822(3)	0.19547(3)	0.19109(3)
	NSIT	0.19518(9)	0.20531(11)	0.20747(17)

Table 5: Example 3: relative errors for NSIWT and NSIFT with the nonstationary α_n in (8.1).

	NSIFT	NSIWT	NSIT
Error	0.19335(10)	0.18765(8)	0.19518(9)

Table 6: Example 3: relative restorations errors for NSIFT and NSIWT with parameters in (8.3) and NSIT with $\alpha_n = 0.01 \cdot 0.7^n$.



(a) (b) (c)

Figure 4: Example 3 – "blur" reconstructions: (a) NSIFT and (b) NSIWT with parameters in (8.3), (c) NSIT with $\alpha_n = 0.01 \cdot 0.7^n$.

Acknowledgement

The authors warmly thank L. Reichel for illuminating discussions. This work is partly supported by PRIN 2012 N. 2012MTE38N for the first three authors, while the work of the fourth author is partly supported by the Program 'Becoming the Number One – Sweden (2014)' of the Knut and Alice Wallenberg Foundation.

References

[1] M Brill and E Schock. Iterative solution of ill-posed problems-a survey. *Model optimization in exploration geophysics, ed. A. Vogel, Vieweg, Braunschweig*, 1987.

[2] Hansen P C. Regularization tools: a Matlab package for analysis and solution of discrete ill-posed problems. *Numer. Algorithms*, 6(1-2):1–35, 1994.

[3] Hansen P C. *Rank-deficient and discrete ill-posed problems.* SIAM Monographs on Mathematical Modeling and Computation. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1998. Numerical aspects of linear inversion.

[4] Hansen P C, Nagy J G, and O'Leary D P. *Deblurring images*, volume 3 of *Fundamentals of Algorithms.* Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2006. Matrices, spectra, and filtering.

[5] Marco Donatelli and Martin Hanke. Fast nonstationary preconditioned iterative methods for ill-posed problems, with application to image deblurring. *Inverse Problems*, 29(9):095008, 2013.

[6] Klann E and Ramlau R. Regularization by fractional filter methods and data smoothing. *Inverse Problems*, 24(2):025018, 2008.

[7] Daniel Gert, Esther Klann, Ronny Ramlau, and Lothar Reichel. On fractional tikhonov regularization. *private notes*, 2014.

[8] Martin Hanke and Charles W Groetsch. Nonstationary iterated tikhonov regularization. *Journal of Optimization Theory and Applications*, 98(1):37–53, 1998.

[9] Michiel E Hochstenbach and Lothar Reichel. Fractional tikhonov regularization for linear discrete ill-posed problems. *BIT Numerical Mathematics*, 51(1):197–215, 2011.

[10] AK Louis. Inverse und schlecht gestellte probleme, teubner, stuttgart, 1989. *MR 90g*, 65075, 1989.

[11] Donatelli M. On nondecreasing sequences of regularization parameters for nonstationary iterated tikhonov. *Numer. Algorithms*, 40(4):651–668., 2012.

[12] Hanke M and Hansen P C. Regularization methods for large-scale problems. *Surveys Math. Indust.*, 3(4):253–315, 1993.

[13] Walter Rudin. *Real and complex analysis.* Tata McGraw-Hill Education, 1987.

[14] Walter Rudin. Functional analysis. international series in pure and applied mathematics, 1991.

[15] Engl H W, Hanke M, and Neubauer A. *Regularization of inverse problems*, volume 375. Springer, 1996.

[16] Groetsch C W. *The theory of Tikhonov regularization for Fredholm equations of the first kind*, volume 105 of *Research Notes in Mathematics.* Pitman (Advanced Publishing Program), Boston, MA, 1984.

Appendix A

Lemma 22
Proof. Obviously, both the series converge or diverge simultaneously due to the Asymptotic Comparison test. If they converge, the thesis follows trivially. On the contrary, if they both diverge then we conclude by observing that $\sum_{k=N}^{n} t_k / \sum_{k=1}^{n} t_k$ is a monotonic increasing sequence bounded from above by 1. Indeed, if we set

$$A_n := \sum_{k=N}^{n} t_k, \qquad B_n := \sum_{k=1}^{n} t_k,$$

for every $n \geq N$ and for every $x \geq 0$ the function

$$h_n(x) = \frac{A_n + x}{B_n + x}$$

is monotone increasing with $h_n(x) \leq 1$. Then $A_{n+1}/B_{n+1} \geq A_n/B_n$ for every n and it is easy to see that $\sup_n\{A_n/B_n\} = 1$. □

Lemma 23 *Proof.* If $\lim_{k\to\infty} t_k = t \in (0, \infty]$, then

$$\frac{1}{t_k} \sim \left(\frac{1}{\lambda} + \frac{1}{t}\right) \frac{1}{1 + t_k}, \tag{8.4}$$

where $1/t = 0$ if $t = \infty$. Therefore, from the Asymptotic Comparison test for series, both series converge or diverge simultaneously. When they converge the thesis follows trivially. Assume then that the series diverge. Without loss of generality and for the sake of clarity we will prove the statement for $\lambda = 1$. If we set

$$X_n := \frac{\sum_{k=1}^{n} \frac{1}{t_k}}{\sum_{k=1}^{n} \frac{1}{1+t_k}},$$

we want to show that the limit of X_n exists finite and, moreover, that is $\lim_{n\to\infty} X_n = 1 + 1/t$. Indeed, for any fixed $\epsilon > 0$ there exists N_ϵ^1 such that for any $k \geq N_\epsilon^1$ it holds that

$$\frac{1}{t_k} < \left(1 + \frac{1}{t} + \frac{\epsilon}{2}\right) \frac{1}{1 + t_k}, \tag{8.5}$$

and for any fixed ϵ and N_ϵ^1, there exists N_ϵ^2 such that for every $n \geq N_\epsilon^2$ it holds that

$$\frac{\sum_{k=1}^{N_\epsilon^1} \frac{1}{t_k}}{\sum_{k=1}^{n} \frac{1}{1+t_k}} < \frac{\epsilon}{2}. \tag{8.6}$$

Hence, for any $n \geq \max\{N_\epsilon^1, N_\epsilon^2\}$, thanks to (8.5) and (8.6), we have that

$$X_n = \frac{\sum_{k=1}^{n} \frac{1}{t_k}}{\sum_{k=1}^{n} \frac{1}{1+t_k}} < \frac{\sum_{k=1}^{N_\epsilon^1} \frac{1}{t_k}}{\sum_{k=1}^{n} \frac{1}{1+t_k}} + \left(1 + \frac{1}{t} + \frac{\epsilon}{2}\right) \frac{\sum_{k=N_\epsilon^1+1}^{n} \frac{1}{1+t_k}}{\sum_{k=1}^{n} \frac{1}{1+t_k}} < \frac{\epsilon}{2} + 1 + \frac{1}{t} + \frac{\epsilon}{2} = 1 + \frac{1}{t} + \epsilon.$$

On the other hand, there exists N_ϵ^3 such that for every $k \geq N_\epsilon^3$ it holds

$$\frac{1}{t_k} > \left(1 + \frac{1}{t} - \frac{\epsilon}{2}\right) \frac{1}{1+t_k}, \tag{8.7}$$

and, by Lemma 22, for any fixed N_ϵ^3 and for any fixed $\delta < \frac{\epsilon}{2}(1 + \frac{1}{t} - \frac{\epsilon}{2})^{-1}$, there exists N_ϵ^4 such that for every $n \geq N_\epsilon^4$ it holds

$$\frac{\sum_{k=N_\epsilon^3+1}^{n} \frac{1}{1+t_k}}{\sum_{k=1}^{n} \frac{1}{1+t_k}} > (1 - \delta)\,. \tag{8.8}$$

Hence, fo any $n \geq \max\{N_\epsilon^3, N_\epsilon^4\}$, thanks to (8.7) and (8.8), we have that

$$X_n = \frac{\sum_{k=1}^{n} \frac{1}{t_k}}{\sum_{k=1}^{n} \frac{1}{1+t_k}} > \frac{\sum_{k=1}^{N_\epsilon^1} \frac{1}{t_k}}{\sum_{k=1}^{n} \frac{1}{1+t_k}} + \left(1 + \frac{1}{t} - \frac{\epsilon}{2}\right) \frac{\sum_{k=N_\epsilon^1+1}^{n} \frac{1}{1+t_k}}{\sum_{k=1}^{n} \frac{1}{1+t_k}} > \left(1 + \frac{1}{t} - \frac{\epsilon}{2}\right)(1-\delta) > 1 + \frac{1}{t} - \epsilon.$$

Then, choosing $n \geq \max\{N_\epsilon^i : i = 1, 2, 3, 4\}$, the proof is concluded. □