ON GENERALIZED ITERATED TIKHONOV REGULARIZATION
WITH OPERATOR-DEPENDENT SEMINORMS[*]

DAVIDE BIANCHI[†] AND MARCO DONATELLI[†]

Abstract. We investigate the recently introduced Tikhonov regularization filters with penalty terms having seminorms that depend on the operator itself. Exploiting the singular value decomposition of the operator, we provide optimal order conditions, smoothing properties, and a general condition (with a minor condition of the seminorm) for the saturation level. Moreover, we introduce and analyze both stationary and nonstationary iterative counterparts of the generalized Tikhonov method with operator-dependent seminorms. We establish their convergence rate under conditions affecting only the iteration parameters, proving that they overcome the saturation result. Finally, some selected numerical results confirm the effectiveness of the proposed regularization filters.

Key words. ill-posed problems, fractional Tikhonov regularization, iterated Tikhonov, filter functions

AMS subject classifications. 65F22, 47A52, 65R32

1. Introduction. We consider an equation of the form

$$(1.1) \qquad Kx = y,$$

where $K : \mathcal{X} \to \mathcal{Y}$ is a compact linear operator between generic Hilbert spaces $\mathcal{X}$ and $\mathcal{Y}$ over $\mathbb{R}$ or $\mathbb{C}$. Let $\mathrm{Rg}(K)$ be the range of K. We assume $y \in \mathrm{Rg}(K)$, i.e., that the problem (1.1) has a solution $x^\dagger = K^\dagger y$ of minimal norm. Here $K^\dagger$ denotes the (Moore-Penrose) pseudo 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 [12].

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

$$(1.2) \qquad y^\delta = y + \eta \qquad \text{and} \qquad \|\eta\| \leq \delta,$$

where η is called the *noise vector* and δ is called the *noise level*. Since $K^\dagger y^\delta$ is not necessarily close to $x^\dagger$, we approximate $x^\dagger$ by $x_\alpha^\delta := R_\alpha y^\delta$, where $\{R_\alpha\}$ is a family of continuous operators depending on a parameter α [7]. In that sense, a class of such regularizers are the so-called filter-based methods that exploit the singular value expansion (s.v.e.) of the operator K by acting on its spectrum in order to diminish the effect of the noise on the reconstructed solution. Indeed, if we indicate by $(\sigma_m; v_m, u_m)_{m \in \mathbb{N}}$ the s.v.e of K, then we can express R_α in the following way:

$$R_\alpha y = \sum_{m=1}^{+\infty} F_\alpha(\sigma_m) \sigma_m^{-1} \langle y, u_m \rangle v_m,$$

for every y that belongs to the domain of $K^\dagger$ and where $\{F_\alpha\}_{\alpha \in \mathbb{R}_+}$ is a family of functions (filters) that, under suitable conditions, compensate the decay to zero of the sequence of singular values $\{\sigma_m\}$ of the compact operator K; see, e.g., [20]. Namely, we can write

$$(1.3) \qquad x_\alpha^\delta = R_\alpha y^\delta = x^\dagger - e_a + e_n,$$

___

[†]Dipartimento di Scienza e Alta Tecnologia, Università dell'Insubria, via Valleggio 11, 22100 Como, Italy (`{d.bianchi9, marco.donatelli}@uninsubria.it`).

where

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

is the *approximation* error and

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

is the *noise* error. The role of the filter function F_α is then to mediate between the approximation error and the noise error.

A classical example is Tikhonov regularization with the filter $F_\alpha(\sigma) = \frac{\sigma^2}{\sigma^2 + \alpha}$. It is well known that the Tikhonov filter is affected by a couple of undesirable properties such as an *oversmoothing* effect on the regularized solution x_α^δ and a *saturation* of the convergence rate. Indeed, if we consider for example the compact embedding operator $J_s : \mathrm{H}^s([0, 2\pi]) \hookrightarrow \mathrm{L}^2([0, 2\pi])$ of the fractional Sobolev space $\mathrm{H}^s([0, 2\pi])$ onto $\mathrm{L}^2([0, 2\pi])$ with $s > 0$, then the regularized solution x_α^δ belongs to the fractional Sobolev space H^{2s}, i.e., x_α^δ lies in a more regular space than the true solution. Furthermore, Tikhonov regularization under suitable a-priori assumptions and an a-priori choice rule, $\alpha = \alpha(\delta) \sim c \cdot \delta^{2/3}$, is of optimal order, and the best possible convergence rate obtainable is

$$\|x_\alpha^\delta - x^\dagger\| = O(\delta^{\frac{2}{3}}).$$

On the other hand, let Q be the orthogonal projector onto $\overline{\mathrm{Rg}(K)}$. If

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

then $x^\dagger = 0$ as long as $\mathrm{Rg}(K)$ is not closed. This shows that Tikhonov regularization for an ill-posed problem with a compact operator never yields a convergence rate that is faster than $O(\delta^{\frac{2}{3}})$ because the method saturates at this rate; see [10, 7].

In the last years, new types of Tikhonov-based regularization methods were studied in [18] and [15] under the name of Fractional or Weighted Tikhonov and in [17, 19] in order to dampen the oversmoothing effect on the regularized solution of classic Tikhonov and to exploit the information carried by the spectrum of the operator. Special attention was devoted to Fractional Tikhonov regularization studied and extended in [9, 1, 15], while for Hermitian problems, the fractional approach was combined with Lavrentiev regularization; see [21, 14].

In this paper we present a generalization of the filters in [15, 17]. We show that the aforementioned methods can be seen as a generalized Tikhonov method of the form

$$(1.4) \qquad R_{\alpha, f} y := \operatorname{argmin}_{x \in \mathcal{X}} \left\{ \|Kx - y\|^2 + \alpha \|x\|^2_{\sqrt{f(K^*K)}} \right\},$$

or equivalently,

$$(1.5) \qquad R_{\alpha, f} y = \left(K^*K + \alpha f\left(K^*K\right) \right)^{-1} K^* y,$$

where $f : [0, \sigma_1] \to \mathbb{R}$ is a bounded measurable function and the associated filter function has the structure

$$(1.6) \qquad F_{\alpha, f}(\sigma) = \frac{\sigma^2}{\sigma^2 + \alpha f(\sigma^2)}.$$

We investigate the regularization and convergence properties of the filter, and we study conditions under which the regularization of the embedding operator $j_s : \mathrm{H}^s([0, 2\pi]) \hookrightarrow \mathrm{L}^2([0, 2\pi])$ recovers the regularized solution in H^t with $t \leq 2s$, i.e., such that the oversmoothing effect of the Tikhonov filter is reduced and more features of the true solution can be recovered. Moreover, a saturation result for the regularization method (1.4) is proposed. To overcome such saturation, we apply the iterative approach as it was done for the classical iterated Tikhonov method [12]. Employing the techniques in [1], we introduce the iterative nonstationary variant for the regularization method (1.4) with the function f defined according to the proposal in [17] and give sufficient conditions for convergence and stability. Nonstationary versions of Tikhonov methods can provide better restorations, see, e.g., [16], as also confirmed by our numerical results, and in particular let us avoid the nontrivial problem of finding the optimal parameter α.

Iterated Tikhonov with a general penalty term has recently been investigated in [4], where both convergence and regularization properties are proved. Here, we consider general penalty terms that are particular functions of the operator, and we prove stronger results. For instance, we provide the convergence rate in the noise-free case and the regularization property in the noisy case under a weaker condition. Moreover, this paper is to be considered an evolution and integration of [1], providing new results and a generalization of some of the techniques introduced in that paper.

The paper is organized as follow. In Section 2 we set the basic notations and preliminaries. In Section 3 we introduce the filter functions which we study and provide order optimality results. In Sections 4 and 5 we investigate the smoothing properties of these methods and their saturation levels, respectively. In Section 6 we present their iterated versions and provide convergence results. In Section 7 we give selected examples which confirm the theoretical analysis, while Section 8 is devoted to concluding remarks.

2. Preliminaries. In this section we recall some classical results on regularization methods that will be used in our subsequent analysis. The main concern is to prescribe sufficient conditions for which a family of approximate solutions converges uniformly to the exact true solution with respect to the noise level and with the best possible upper bound on the rate of convergence. All the propositions and theorems that are not new will include a reference to the paper where they first appeared. For a detailed description of the following results, see [7, 20].

DEFINITION 2.1 (Generalized Inverse). *We define the* generalized inverse *of a compact linear operator* $K : \mathcal{X} \to \mathcal{Y}$ *as* $K^\dagger : \mathrm{Dom}(K^\dagger) \subseteq \mathcal{Y} \to \mathcal{X}$,

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

where

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

As discussed in the introduction, we consider the problem (1.1) where only an approximation y^δ of y satisfying (1.2) is available. Therefore, $x^\dagger = K^\dagger y$ with $y \in \mathrm{Dom}(K^\dagger)$ 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 $\{R_\alpha\}$ of neighboring well-posed problems [7], where R_α should be continuous for every α and $R_\alpha \to K^\dagger$ pointwise for $\alpha \to 0$. These properties are formalized in the next definition.

DEFINITION 2.2. *By a* regularization method *for $K^\dagger$ we mean 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 $y \in \mathrm{Dom}(K^\dagger)$ there exists a mapping $\alpha : \mathbb{R}_+ \times \mathcal{Y} \to (0, \alpha_0) \in \mathbb{R}$, (i.e., a choice rule) with $\alpha = \alpha(\delta, y^\delta)$, 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 distinguish them by means of a subscript.

The special regularization methods which we are interested in are filtering-based methods largely investigated in [20].

PROPOSITION 2.3 ([20]). *Let $K : \mathcal{X} \to \mathcal{Y}$ be a compact linear operator. Let $\sigma(K)$ be the closure of $\bigcup_{m=1}^\infty \{\sigma_m\}$, and let $K^\dagger$ be 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*

$$(2.1) \qquad R_\alpha y := \sum_{m : \sigma_m > 0} F_\alpha(\sigma_m) \sigma_m^{-1} \langle y, u_m \rangle v_m,$$

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

$$(2.2a) \qquad \sup_{m : \sigma_m > 0} |F_\alpha(\sigma_m) \sigma_m^{-1}| = c(\alpha) < \infty,$$

$$(2.2b) \qquad |F_\alpha(\sigma_m)| \leq c < \infty, \qquad \text{where } c \text{ does not depend on } (\alpha, m),$$

$$(2.2c) \qquad \lim_{\alpha \to 0} F_\alpha(\sigma_m) = 1 \text{ pointwise in } \sigma_m.$$

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

For ease of notation we introduce the following notations

$$(2.3) \qquad x_\alpha := R_\alpha y, \qquad y \in \mathrm{Dom}(K^\dagger),$$

$$(2.4) \qquad x_\alpha^\delta := R_\alpha y^\delta, \qquad y^\delta \in \mathcal{Y}.$$

A choice rule $\alpha(\delta, y^\delta)$ that depends only on the noise level δ, i.e., $\alpha(\delta, y^\delta) = \alpha(\delta)$, is called *a-priori* choice rule. Since there can be no uniform convergence for any regularization method if $\mathrm{Rg}(K)$ is not closed, see [7, Proposition 3.11], convergence rates can only be proven on subsets of $\mathcal{X}$, namely under *a-priori assumptions* in terms of the exact solution.

We thus report hereafter the definition of optimal order under a-priori assumptions.

DEFINITION 2.4 (Optimal order under an a-priori assumption). *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*

$$\Delta(\delta, \mathcal{X}_{\nu,\rho}, R_\alpha) \le c \cdot \delta^{\frac{\nu}{\nu+1}} \rho^{\frac{1}{\nu+1}}, \tag{2.5}$$

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

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

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

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

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

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

Roughly speaking, depending on the space $\mathcal{X}$, $\mathcal{X}_{\nu,\rho}$ describes the regularity of the exact solution $x^\dagger$ in terms of subdomains of $\mathcal{X}$.

Finally, we are ready to state the main result of this section and one of the most important tools that we use for the study of the upcoming new filter methods. The next theorem provides sufficient conditions for a regularization method to be of optimal order under an a-priori assumption, i.e., it provides sufficient conditions in order to achieve the best upper bound on the convergence rate with respect to the noise level.

THEOREM 2.5 ([20]). *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*

$$\sup_{0 < \sigma \le \sigma_1} |F_\alpha(\sigma)\sigma^{-1}| \le c \cdot \alpha^{-\beta}, \tag{2.6a}$$

$$\sup_{0 \le \sigma \le \sigma_1} |(1 - F_\alpha(\sigma))\sigma^\nu| \le c_\nu \cdot \alpha^{\beta\nu}, \tag{2.6b}$$

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

$$\alpha = \alpha(\delta, \rho) = \hat{c} \cdot \left(\frac{\delta}{\rho} \right)^{\frac{1}{\beta(\nu+1)}}, \qquad 0 < \hat{c} = \left(\frac{c}{\nu c_\nu} \right)^{\frac{1}{\beta(\nu+1)}}.$$

3. Two generalized Tikhonov methods and a mixing approach. In this section we discuss two recent types of regularization methods which generalize the classical Tikhonov method and which were first introduced and studied in [18] and [17], plus a mixing method with the intent to merge the features of both the preceding methods. We use the notation $F_{\alpha,\cdot}$ to indicate the new filters, where $\cdot$ will be replaced by the extra parameter introduced by the respective method. Every method is studied separately to avoid confusion and misunderstandings.

DEFINITION 3.1 ([15]). *We define the* weighted-I Tikhonov *method as 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

$$(3.1) \qquad F_{\alpha,r}(\sigma) = \frac{\sigma^{r+1}}{\sigma^{r+1} + \alpha} = \frac{\sigma^2}{\sigma^2 + \alpha\sigma^{1-r}},$$

for $\alpha > 0$ and $r \geq 0$. For $r = 1$ the classic Tikhonov filter is recovered.

According to (2.3) and (2.4), we fix the following notation

$$x_{\alpha,r} := R_{\alpha,r}y, \qquad y \in \mathrm{Dom}(K^\dagger),$$
$$x_{\alpha,r}^\delta := R_{\alpha,r}y^\delta, \qquad y^\delta \in \mathcal{Y}.$$

DEFINITION 3.2 ([17]). *We define the* weighted-II Tikhonov *method as the filter-based method*

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

where the filter function is

$$(3.2) \qquad F_{\alpha,j}(\sigma) = \frac{\sigma^2}{\sigma^2 + \alpha\left[1 - \left(\frac{\sigma}{\sigma_1}\right)^2\right]^j},$$

for $\alpha > 0$ and $j \in \mathbb{N}$. For $j = 0$ the classic Tikhonov filter is recovered.

REMARK 3.3. With reference to (1.3), let us observe that for large j, the weighted-II filter $F_{\alpha,j}(\sigma)$ is almost 1 when σ belongs to the signal subspace and is almost the standard Tikhonov filter $F_\alpha(\sigma) = \frac{\sigma^2}{\sigma^2 + \alpha}$ when σ belongs to the noise subspace. The idea is that in the signal subspace, i.e., when $\sigma \sim \sigma_1$, where the noise error norm is controlled, the regularization is minimal, avoiding as much as possible the approximation error, while the action of the filter function is focused on the noise subspace, i.e., when $\sigma \sim 0$. Roughly speaking, the weighted-II filter acts like a switch for the regularization to take place. Like above, we fix the following notation

$$x_{\alpha,j} := R_{\alpha,j}y, \qquad y \in \mathrm{Dom}(K^\dagger),$$
$$x_{\alpha,j}^\delta := R_{\alpha,j}y^\delta, \qquad y^\delta \in \mathcal{Y}.$$

Given an operator W on any Hilbert space, if we consider the seminorm $\|\cdot\|_W$ induced by W, i.e., $\|x\|_W^2 := \langle Wx, Wx\rangle$, then the weighted-I Tikhonov method can also be defined as the unique minimizer of the functional

$$(3.3) \qquad R_{\alpha,r}y := \mathrm{argmin}_{x\in\mathcal{X}}\left\{\|Kx - y\|^2 + \alpha\|x\|_W^2\right\},$$

where the seminorm $\|\cdot\|_W$ is induced by the operator $W := (K^*K)^{\frac{1-r}{4}} : \mathcal{X} \to \mathcal{X}$. For $r > 1$, W has to be interpreted as the Moore-Penrose pseudo inverse . Looking for a stationary point in equation (3.3), we have that

$$0 = \nabla\left(\|Kx - y\|^2 + \alpha\|x\|_W^2\right)$$
$$= \nabla\left(\langle Kx, Kx\rangle - 2\langle Kx, y\rangle + \langle y, y\rangle + \alpha\langle(K^*K)^{\frac{1-r}{4}}x, (K^*K)^{\frac{1-r}{4}}x\rangle\right)$$
$$= \nabla\left(\langle x, K^*Kx\rangle - 2\langle x, K^*y\rangle + \alpha\langle x, (K^*K)^{\frac{1-r}{2}}x\rangle\right)$$
$$= 2K^*Kx - 2K^*y + 2\alpha(K^*K)^{\frac{1-r}{2}}x,$$

from which we deduce the following expression for the operator $R_{\alpha,r}$,

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

In the same way, the weighted-II Tikhonov method can be defined as the unique minimizer of the functional

$$R_{\alpha,j}y := \operatorname{argmin}_{x\in\mathcal{X}}\left\{\|Kx - y\|^2 + \alpha\|x\|_B^2\right\},$$

where $B := \left(I - \frac{K^*K}{\|K^*K\|}\right)^{j/2} : \mathcal{X} \to \mathcal{X}$, and by similar calculations as above, we can deduce that

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

Both these methods can be classified in the more general context of weighted generalized inverse methods, namely as

(3.4) $$R_\alpha y := \operatorname{argmin}_{x\in\mathcal{X}}\left\{\|Kx - y\|^2 + \alpha\|x\|_\Lambda^2\right\},$$

or

(3.5) $$R_\alpha y = \left[K^*K + \alpha\Lambda^*\Lambda\right]^{-1}K^*y,$$

where Λ is a suitable operator. We do not go into details; for references, see [7, Chapter 8]. We just observe that if $\Lambda^*\Lambda$ and K^*K commute, then, indicating by $(\lambda_n; v_n, u_n)_{n\in\mathbb{N}}$ the s.v.e. of Λ, the operator (3.5) can be expressed as

(3.6) $$R_\alpha y := \sum_{m:\,\sigma_m>0} F_\alpha(\sigma_m,\lambda_m)\sigma_m^{-1}\langle y, u_m\rangle v_m, \qquad \text{with } F_\alpha(\sigma,\lambda) = \frac{\sigma^2}{\sigma^2 + \alpha\lambda^2}.$$

Now, let $f : [0, \sigma_1] \to [0, \infty)$ be a Borel-measurable function and consider the operator $f(K^*K)$, which commutes with K^*K. From equations (3.6), (3.1), and (3.2), it is clear that both the weighted-I and weighted-II filter methods are of the form (3.4) with $\Lambda = \sqrt{f(K^*K)}$ and where

$$f(\sigma^2) = \sigma^{1-r} \quad \text{and} \quad f(\sigma^2) = \left(1 - \left(\frac{\sigma}{\sigma_1}\right)^2\right)^j,$$

respectively. This is the reason that motivated us to rename the original method of Hochstenbach and Reichel, which appeared in [15], into *weighted-I Tikhonov method* and subsequently to rename the method of Huckle and Sedlacek appearing in [17] into *weighted-II Tikhonov method*. In this way it is easier to distinguish them from the *fractional Tikhonov method* of Klann and Ramlau in [18].

We can now introduce a new mixed method that makes use of both of the weighted-I and weighted-II methods by combining their filter functions.

DEFINITION 3.4. *Fixing $F_{\alpha,r,j}$ such that*

$$F_{\alpha,r,j}(\sigma) := \frac{\sigma^2}{\sigma^2 + \alpha f(\sigma^2)} = \frac{\sigma^2}{\sigma^2 + \alpha f_1(\sigma^2)f_2(\sigma^2)},$$

where

$$f_1(\sigma^2) = \sigma^{1-r}, \qquad f_2(\sigma^2) = \left[1 - \left(\frac{\sigma}{\sigma_1}\right)^2\right]^j,$$

we define the mixed *method as the filter-based method*

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

Let us notice that for $r = 1$ and $j = 0$ we recover the standard Tikhonov method.

As above, we use the notation

$$x_{\alpha,r,j} := R_{\alpha,r,j}y, \qquad y \in \mathrm{Dom}(K^\dagger),$$
$$x_{\alpha,r,j}^\delta := R_{\alpha,r,j}y^\delta, \qquad y^\delta \in \mathcal{Y}.$$

The idea behind this method is to exploit the features of the weighted-II filter function, as observed in Remark 3.3, and use it as a switch for the regularization. Roughly speaking,

$$F_{\alpha,r,j}(\sigma) \sim 1 \text{ when } \sigma \sim \sigma_1, \qquad F_{\alpha,r,j}(\sigma) \sim F_{\alpha,r}(\sigma) \text{ when } \sigma \sim 0.$$

This regularization method satisfies all the nice properties of the weighted-I method that we are going to present soon after (see Propositions 3.5, 4.1, and 5.1): it is of optimal order, it saturates at the rate of $O\left(\delta^{\frac{r+1}{r+2}}\right)$, it undersmooths the reconstructed solution for $0 < r < 1$, and it oversmooths the reconstructed solution for $r \geq 1$. For brevity, we skip most of the details of the proofs of statements regarding the mixed method. Indeed they can be easily recovered adapting and combining the techniques used in Propositions 3.5, 3.6, 4.2, and 5.1.

Proceeding further with the analysis, the optimal order of the weighted-I Tikhonov regularization was proved in [9]. The following proposition summarizes this result, highlighting the dependence on r of ν, and provides a converse result.

PROPOSITION 3.5 ([15]). *Let K be a compact linear operator with infinite-dimensional range. For every given $r \geq 0$, the weighted-I Tikhonov method $R_{\alpha,r}$ is a regularization method of optimal order under the a-priori assumption $x^\dagger \in \mathcal{X}_{\nu,\rho}$ with $0 < \nu \leq r + 1$. The best possible rate of convergence with respect to δ is $\|x^\dagger - x_{\alpha,r}^\delta\| = O(\delta^{\frac{r+1}{r+2}})$, which 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 $x^\dagger \in \mathcal{X}_{r+1}$.*

The next proposition provides a proof for the optimal order of the weighted-II Tikhonov regularization, which instead fails to have a converse result, namely it is not possible to infer any regularity of $x^\dagger$ from the convergence rate of $\|x^\dagger - x_{\alpha,r}\|$.

PROPOSITION 3.6. *Let K be a compact linear operator with infinite-dimensional range. For every given integer $j \geq 0$, the weighted-II Tikhonov method $R_{\alpha,j}$ is a regularization method of optimal order under the a-priori assumption $x^\dagger \in \mathcal{X}_{\nu,\rho}$ with $0 < \nu \leq 2$. The best possible rate of convergence with respect to δ is $\|x^\dagger - x_{\alpha,j}^\delta\| = O\left(\delta^{\frac{2}{3}}\right)$, which is obtained for $\alpha = \left(\frac{\delta}{\rho}\right)^{\frac{2}{\nu+1}}$ with $\nu = 2$.*

Proof. The validity of (2.2a), (2.2b), and (2.2c) are trivial to verify, having as a consequence that the weighted-II Tikhonov is a regularization filter method. Without loss of generality, we suppose that $\sigma_1 = 1$. The left-hand side of condition (2.6b) takes the form

$$\sup_{0 \leq \sigma \leq 1} \left|\frac{\alpha\left(1 - \sigma^2\right)^j \sigma^\nu}{\sigma^2 + \alpha\left(1 - \sigma^2\right)^j}\right|,$$
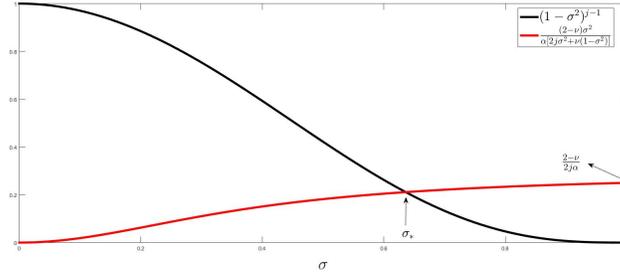
FIGURE 3.1. *Graphic solution for σ_*.*

which is bounded from above by

$$\sup_{0 \le \sigma \le 1} |\alpha g(\sigma)|, \qquad g(\sigma) = \frac{\sigma^\nu}{\sigma^2 + \alpha \left(1 - \sigma^2\right)^j}.$$

Let us study the function $g(\sigma)$. By taking the derivative, we observe that σ^* is a maximal point of g if and only if $\nu \in [0, 2]$ and σ^* satisfies the equation

$$(1 - \sigma^2)^{j-1} = \frac{(2 - \nu)\sigma^2}{\alpha \left[2j\sigma^2 + \nu \left(1 - \sigma^2\right)\right]}.$$

It is not difficult to see that there exists only one such $\sigma_* \in [0, 1]$; see Figure 3.1. Indeed, the functions

$$\psi_1(\sigma) = (1 - \sigma^2)^{j-1}, \qquad \psi_2(\sigma) = \frac{(2 - \nu)\sigma^2}{\alpha \left[2j\sigma^2 + \nu \left(1 - \sigma^2\right)\right]}$$

are such that $\psi_1' \le 0$ and $\psi_2' > 0$ on $(0, 1)$. Therefore, if we define

$$\phi(\sigma) = \psi_2(\sigma) - \psi_1(\sigma),$$

and fixing $\alpha_0 < 2/(2 - \nu)$ in the case $j = 1$, it holds that

$$\phi(0) = -1, \qquad 0 < \phi(1) = \begin{cases} \frac{2-\nu}{2\alpha} - 1 & \text{if } j = 1 \\ \frac{2-\nu}{2j\alpha} & \text{if } j > 1 \end{cases}, \qquad \phi' > 0 \text{ on } (0, 1).$$

By standard calculus arguments, there exists only one σ^* with $\phi(\sigma^*) = 0$, namely, the value $\sigma_* = \sqrt{\alpha} h(\sigma)$, where

$$h(\sigma) = \left\{ \frac{(1 - \sigma^2)^{j-1} \left[2j\sigma^2 + \nu \left(1 - \sigma^2\right)\right]}{2 - \nu} \right\}^{1/2}.$$

If we let $\alpha \in (0, \alpha_0)$ with $\alpha_0 < \infty$, then necessarily $\sigma_* \in (0, \lambda_{\alpha_0, j})$, where $\lambda_{\alpha_0, j} < 1$. Since $h(\sigma_*) = 0$ if and only if $\sigma_* = 1$, $h(\sigma_*)$ is uniformly bounded away from 0, i.e., $h(\sigma_*) \in [\hat{c}, 1]$ with $\hat{c} = \hat{c}(\alpha_0, j, \nu) > 0$ and independent of α. Henceforth, we can write

$$\sigma_* = c_{\alpha_0, j, \nu} \sqrt{\alpha} = c\sqrt{\alpha}$$

with $0 < c \in [\hat{c}, 1]$, and then we have

$$
(3.7) \quad
\begin{aligned}
\sup_{0 \leq \sigma \leq 1} |(1 - F_{\alpha,j}(\sigma)) \sigma^\nu| &= \sup_{0 \leq \sigma \leq 1} \left| \frac{\alpha \left(1 - \sigma^2\right)^j \sigma^\nu}{\sigma^2 + \alpha \left(1 - \sigma^2\right)^j} \right| \\
&\leq \sup_{0 \leq \sigma \leq 1} |\alpha g(\sigma)| = \alpha g(\sigma_*) = \alpha g(c\sqrt{\alpha}) = \frac{c^\nu \alpha^{\nu/2}}{c^2 + (1 - c^2 \alpha)^j} \leq c^{\nu-2} \alpha^{\nu/2},
\end{aligned}
$$

which is (2.6b) with $\beta = 1/2$. The validity of (2.6a) is shown again by studying the function g with fixed $\nu = 1$. Therefore, from Theorem 2.5, as long as $0 < \nu \leq 2$, if $x^\dagger \in \mathcal{X}_{\nu,\rho}$, the method achieves order optimality (2.5), and the best possible rate of convergence obtainable with respect to δ is $O(\delta^{\frac{\nu}{\nu+1}})$ for $\nu = 2$.   □

REMARK 3.7. Observe that the optimal order for the weighted-II Tikhonov is independent of the auxiliary parameter j.

COROLLARY 3.8. *Let K be a compact linear operator with infinite-dimensional range. For every given $r \geq 0, j \in \mathbb{N}$, the mixed method $R_{\alpha,r,j}$ is a regularization method of optimal order under the a-priori assumption $x^\dagger \in \mathcal{X}_{\nu,\rho}$ with $0 < \nu \leq r + 1$. The best possible rate of convergence with respect to δ is $\|x^\dagger - x^\delta_{\alpha,r,j}\| = O(\delta^{\frac{r+1}{r+2}})$, which is obtained for $\alpha = \left(\frac{\delta}{\rho} \right)^{\frac{r+1}{\nu+1}}$ with $\nu = r + 1$. On the other hand, if $\|x^\dagger - x_{\alpha,r,j}\| = O(\alpha)$, then $x^\dagger \in \mathcal{X}_{r+1}$.*

Proof. Exploiting the structure of $F_{\alpha,r,j}(\sigma) = \frac{\sigma^2}{\sigma^2 + \alpha f_1(\sigma^2) f_2(\sigma^2)}$ with f_1 and f_2 as in Definition 3.4, the proof can be adapted without difficulties combining the techniques used in the proofs of Proposition 3.6 and [1, Proposition 10].   □

4. Smoothing effect. In this section we deal with the *oversmoothing* property that affects the classical Tikhonov regularization method. Indeed, it was observed that the approximate solution is smoother than the true solution, i.e., it lives in a space of higher regularity. We will see that the weighted-I and fractional filters can overcome the oversmoothing effect by a proper choice of their extra regularization parameters. In order to understand this kind of behavior easily, we are going to restrict our study to the fractional Sobolev spaces H^s of one-dimensional functions.

Let $\Omega = [0, 2\pi]$ and let $J_s : H^s(\Omega) \hookrightarrow L^2(\Omega)$ be the embedding operator of the Hilbert space $H^s(\Omega)$ with $s \in (0, \infty)$. J_s is compact with s.v.e. given by

$$
v_m(t) = \left(1 + m^2\right)^{-s/2} e^{imt}, \qquad u_m = e^{imt}, \qquad \sigma_m = \left(1 + m^2\right)^{-s/2}.
$$

Let us consider the following equation

$$
(4.1) \quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad J_s x = y.
$$

Since J_s is compact and therefore ill-conditioned, we regularize the above problem by introducing a family of filter functions,

$$
x^\delta_\alpha(t) = \sum_{m > 0} F_\alpha \left(\sigma_m\right) \sigma_m^{-1} \langle \cdot, u_m \rangle v_m(t).
$$

The true solution $x^\dagger$, i.e., $J_s x^\dagger = y$, belongs to H^s. We consider the regularity of the approximate solution x^δ_α when dealing with general inexact data $y^\delta \in L^2$. The next propositions clarify the H^p spaces in which the approximate solutions live according to the filter method.

PROPOSITION 4.1 ([15]). *For data $y^\delta \in L^2(\Omega)$, the approximate solution $x^\delta_{\alpha,r}$ of the weighted-I Tikhonov filter for equation (4.1) belongs to $H^{s(r+1)}(\Omega)$.*

PROPOSITION 4.2. *For data $y^\delta \in L^2(\Omega)$, the approximate solution $x^\delta_{\alpha,j}$ of the weighted-II Tikhonov filter for the problem (4.1) belongs to $H^{2s}(\Omega)$ for any $j \in \mathbb{N}$.*

Proof. We have that

$$x_{\alpha,j}^{\delta}(t) = \sum_{m>0} \left(\frac{\sigma_m^2}{\sigma_m^2 + \alpha \left[1 - \sigma_m^2\right]^j} \right) \sigma_m^{-1} \langle y^{\delta}, u_m \rangle v_m(t)$$

$$= \sum_{m>0} \left(\frac{(1+m^2)^{-s}}{(1+m^2)^{-s} + \alpha \left[1 - \left(\frac{1+m^2}{2}\right)^{-s}\right]^j} \right) \langle y^{\delta}, u_m \rangle \mathrm{e}^{\mathrm{i}mt}.$$

Then, the Fourier coefficients of $x_{\alpha,j}^{\delta}$ are given by

$$\left(x_{\alpha,j}^{\delta}\right)_m = \left(\frac{(1+m^2)^{-s}}{(1+m^2)^{-s} + \alpha \left[1 - \left(\frac{1+m^2}{2}\right)^{-s}\right]^j} \right) \langle y^{\delta}, u_m \rangle,$$

from which we can calculate the H^p-norm,

$$\|x_{\alpha,j}^{\delta}\|_{\mathrm{H}^p}^2 = \sum_{m>0} (1+m^2)^p \left(\frac{(1+m^2)^{-s}}{(1+m^2)^{-s} + \alpha \left[1 - \left(\frac{1+m^2}{2}\right)^{-s}\right]^j} \right)^2 \left|\langle y^{\delta}, u_m \rangle\right|^2$$

$$= \sum_{m>0} (1+m^2)^{p-2s} \left(\frac{1}{(1+m^2)^{-s} + \alpha \left[1 - \left(\frac{1+m^2}{2}\right)^{-s}\right]^j} \right)^2 \left|\langle y^{\delta}, u_m \rangle\right|^2.$$

Setting $w := (1+m^2)^{-s} \in (0, 2^{-s}]$, it is not difficult to prove that the function

$$\phi(w) = \left(w + \alpha(1 - 2^s w)^j \right)^{-1}$$

is bounded and

$$\min\{2^{-s}; \alpha^{-1}\} \le \phi(w) \le \begin{cases} \alpha^{-1} & \text{if } 0 < \alpha \le (2^s j)^{-1}, \\ \dfrac{1}{\left[1 - (2^s j \alpha)^{-\frac{1}{j-1}}\right] 2^{-s} + \alpha(2^s j \alpha)^{-\frac{j}{j-1}}} & \text{if } \alpha > (2^s j)^{-1}. \end{cases}$$

Therefore, since

$$\frac{1}{(1+m^2)^{-s} + \alpha \left[1 - \left(\frac{1+m^2}{2}\right)^{-s}\right]^j} \le \max_{w \in [0, 2^{-s}]} \phi(w),$$

we can conclude that

$$\|x_{\alpha,j}^{\delta}\|_{\mathrm{H}^p}^2 \le c_{\alpha,j,s} \sum_{m>0} (1+m^2)^{p-2s} \left|\langle y^{\delta}, u_m \rangle\right|^2.$$

The right hand-side of the above inequality is bounded for every data $y^{\delta} \in \mathrm{L}^2(\Omega)$ if $p \le 2s$. $\square$

In Proposition 4.1, for $r = 1$, we recover the classical Tikhonov filter and its *oversmoothing* property, i.e., if the true solution $x^{\dagger} \in \mathrm{H}^s$, then the approximate solution $x_{\alpha}^{\delta} \in \mathrm{H}^{2s}$. Therefore, compared to the standard Tikhonov filter, the weighted-I Tikhonov filter smoothes the approximate solution less for every $0 < r < 1$. We say that it *undersmooths* the approximate solution. The weighted-II Tikhonov instead does not provide any *undersmoothing* effect for any $j \in \mathbb{N}$.

5. Saturation results. The following proposition deals with a saturation property similar to a well-known result for classical Tikhonov, cf. [7, Proposition 5.3]. We extend it to generalized Tikhonov methods of the form (1.4) with the penalty term being measured by the operator-dependent seminorm $\sqrt{f(K^*K)}$, where $f : [0, \sigma_1] \to [0, \infty)$ is a Borel-measurable function such that the corresponding filter function (1.6) satisfies the properties in (2.2).

Again, we fix the following notation

$$\begin{aligned}
x_{\alpha,f} &:= R_{\alpha,f}y, & y &\in \text{Dom}(K^\dagger), \\
x_{\alpha,f}^\delta &:= R_{\alpha,f}y^\delta, & y^\delta &\in \mathcal{Y}.
\end{aligned}$$

PROPOSITION 5.1 (Saturation for weighted Tikhonov regularization). *Let $K : \mathcal{X} \to \mathcal{Y}$ be a compact linear operator with infinite-dimensional range, and let $R_{\alpha,f}$ be the corresponding family of regularization operators as in equation (1.5). Let $\alpha = \alpha(\delta, y^\delta)$ be any parameter choice rule, and let*

$$\tag{5.1} \frac{\sigma^2}{f(\sigma^2)} \sim c\sigma^s \qquad \text{as } \sigma \to 0,$$

with $c, s > 0$. If

$$\tag{5.2} \sup\left\{ \|x_{\alpha,f}^\delta - x^\dagger\| : \|Q(y - y^\delta)\| \leq \delta \right\} = o(\delta^{\frac{s}{s+1}}),$$

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

Proof. Define

$$\begin{aligned}
\delta_m &:= \sigma_m^{s+1}, & y_m^\delta &:= y + \delta_m u_m & \text{so that } \|y - y_m^\delta\| \leq \delta_m, \\
\alpha_m &:= \alpha(\delta_m, y_m^\delta), & x_m &:= x_{\alpha_m,f}, & x_m^\delta &:= x_{\alpha_m,f}^{\delta_m}.
\end{aligned}$$

By the assumption that K has infinite-dimensional range, it follows that $\lim_{m\to\infty} \sigma_m = 0$. According to the expression of $R_{\alpha,f}$ in (1.5), we have

$$\begin{aligned}
x_m^\delta - x^\dagger &= R_{\alpha_m,f}y_m^\delta - x^\dagger = R_{\alpha_m,f}y + \delta_m R_{\alpha_m,f}u_m - x^\dagger \\
&= x_m - x^\dagger + \delta_m F_{\alpha_m,f}(\sigma_m)\sigma_m^{-1}v_m.
\end{aligned}$$

Hence by (3.1), it holds that

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

Since $F_{\alpha,f}$ satisfies (2.2a), we can deduce that f cannot be identically zero in any interval of the form $[0, \lambda]$, and therefore it is possible to divide by $f(\sigma_m^2)$ if we take a suitable subsequence $\{\sigma_{m_n}\} \subseteq \{\sigma_m\}$. Without loss of generality, we assume $\{\sigma_m\} = \{\sigma_{m_n}\}$. Then, we have that

$$\begin{aligned}
\left[\left(\frac{\delta_m\sigma_m}{f(\sigma_m^2)}\right)^{-\frac{1}{2}} \|x_m^\delta - x^\dagger\|\right]^2 &\geq \frac{2}{\frac{\sigma_m^2}{f(\sigma_m^2)} + \alpha_m} \text{Re}\langle x_m - x^\dagger, v_m\rangle + \frac{\delta_m\sigma_m f(\sigma_m^2)}{[\sigma_m^2 + \alpha_m f(\sigma_m^2)]^2} \\
&= \left[\frac{\frac{2f(\sigma_m^2)}{\sigma_m^2}}{1 + \alpha_m \frac{f(\sigma_m^2)}{\sigma_m^2}}\right] \text{Re}\langle x_m - x^\dagger, v_m\rangle + \frac{\delta_m\sigma_m^{-3}f(\sigma_m^2)}{\left[1 + \alpha_m \frac{f(\sigma_m^2)}{\sigma_m^2}\right]^2},
\end{aligned}$$

and passing to the $\limsup$, recalling assumption (5.1) and that $\delta_m = \sigma_m^{s+1}$, we get

(5.3)
$$\limsup_{m\to\infty} \left(\delta_m^{-\frac{s}{s+1}} \|x_m^\delta - x^\dagger\|\right)^2 \geq c \left\{ \limsup_{\sigma_m\to\infty} \left[\frac{2\delta_m^{-\frac{s}{s+1}}}{1 + \alpha_m \delta_m^{-\frac{s}{s+1}}}\right] \operatorname{Re}\langle x_m - x^\dagger, v_m\rangle \right.$$
$$\left. + \liminf_{m\to\infty} \frac{1}{\left[1 + \alpha_m \delta_m^{-\frac{s}{s+1}}\right]^2} \right\},$$

where c is a positive constant.

Thanks to equation (1.5), we have

$$(K^*K + \alpha_m f(K^*K))(x^\dagger - x_m^\delta) = K^*Kx^\dagger + \alpha_m f(K^*K)x^\dagger - K^*y_m^\delta$$
$$= \alpha_m f(K^*K)x^\dagger - \delta_m K^* u_m,$$

thus,

$$\alpha_m x^\dagger = \delta_m [f(K^*K)]^{-1} K^* u_m + \left([f(K^*K)]^{-1} K^*K + \alpha_m I\right)(x^\dagger - x_m^\delta)$$
$$= \frac{\delta_m \sigma_m}{f(\sigma_m^2)} v_m + \left([f(K^*K)]^{-1} K^*K + \alpha_m I\right)(x^\dagger - x_m^\delta).$$

By hypothesis, $\lim_{\sigma\to 0} \sigma^2/f(\sigma^2) = \lim_{\sigma\to 0} c\sigma^s = 0$, and by item (ii) in Definition 2.2, it holds that

$$\alpha_m \leq \sup\left\{\alpha(\delta_m, y^{\delta_m}) : y^{\delta_m} \in \mathcal{Y}, \|y - y^{\delta_m}\| \leq \delta_m\right\} \longrightarrow 0 \quad \text{as } \delta_m \to 0,$$

i.e., $\{\alpha_m\}$ is uniformly bounded. Henceforth,

$$\|[f(K^*K)]^{-1} K^*K + \alpha_m I\| \leq c, \quad \text{for every } m \in \mathbb{N},$$

and then

(5.4)
$$\alpha_m \|x^\dagger\| = O\left(\frac{\delta_m \sigma_m}{f(\sigma_m^2)} + \|x^\dagger - x_m^\delta\|\right).$$

Since $\delta_m = \sigma_m^{s+1}$ and, again from $\lim_{m\to\infty} \sigma_m^2/f(\sigma_m^2) = \lim_{m\to\infty} c\sigma_m^s$, it follows from (5.4) that

$$\alpha_m \|x^\dagger\| \leq c\left(\delta_m^{\frac{2s}{s+1}} + \|x^\dagger - x_m^\delta\|\right).$$

Then, if $x^\dagger \neq 0$,

(5.5)
$$\lim_{m\to\infty} \alpha_m \delta_m^{-\frac{s}{s+1}} = 0$$

because by assumption, $\|x^\dagger - x_m^\delta\| = o\left(\delta_m^{\frac{s}{s+1}}\right)$.

Hence, the second term on the right-hand side of (5.3) tends to 1. Since by assumption the left-hand side of (5.3) tends to 0, we obtain

$$0 \geq c \left\{\limsup_{m\to\infty} \frac{2}{1 + \delta_m^{-\frac{s}{s+1}} \alpha_m} \delta_m^{-\frac{s}{s+1}} \operatorname{Re}\langle x_m - x^\dagger, v_m\rangle + 1\right\}.$$

Now, from (5.2) we have that $\|x_m - x^\dagger\| = o\left(\delta^{\frac{s}{s+1}}\right)$ as well, so that if $x^\dagger \neq 0$, we obtain from (5.5) applied to the preceding inequality the contradiction $0 \geq c > 0$. Hence, $x^\dagger = 0$. ☐

Note that for $f(\sigma^2) \equiv 1$ (classical Tikhonov) the previous proposition gives exactly Proposition 5.3 in [7] with $s = 2$.

For $f(\sigma^2) = \sigma^{1-r}$ and $f(\sigma^2) = \left[1 - \left(\frac{\sigma}{\sigma_1}\right)^2\right]^j$, we recover saturation results for the weighted-I and weighted-II regularization methods, respectively. Indeed,

$$\frac{\sigma^2}{\sigma^{1-r}} \sim \sigma^{r+1}, \qquad \frac{\sigma^2}{\left(1 - \left(\frac{\sigma}{\sigma_1}\right)^2\right)^j} \sim \sigma^2 \quad \text{for } \sigma \to 0.$$

We can state the following corollaries.

COROLLARY 5.2 ([1]). *With the same notation of Proposition 5.1, let $R_{\alpha,r}$ be the family of regularization operators as in Definition 3.1. If*

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

then $x^\dagger = 0$.

Observe that by taking a large value of r, it is possible to overcome the saturation result of classical Tikhonov and obtain a convergence rate arbitrarily close to $O(\delta)$.

COROLLARY 5.3. *With the same notation of Proposition 5.1, let $R_{\alpha,j}$ be the family of regularization operators as in Definition 3.2. If*

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

then $x^\dagger = 0$.

In this case instead, weighted-II Tikhonov saturates at the same level as classical Tikhonov, independently of the choice of the parameter j.

COROLLARY 5.4. *With the same notation of Proposition 5.1, let $R_{\alpha,r,j}$ be the family of regularization operators as in Definition 3.4. If*

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

then $x^\dagger = 0$.

Proof. This follows immediately from

$$\frac{\sigma^2}{\sigma^{1-r}\left(1 - \left(\frac{\sigma}{\sigma_1}\right)^2\right)^j} \sim \sigma^{r+1}. \qquad ☐$$

6. Iterated weighted-II Tikhonov regularization. To bypass the saturation property observed in the previous Section 5, we now propose an iterated regularization method based on the weighted-II Tikhonov filter, see Proposition 6.3, following the same approach adopted in [1]. We first investigate the stationary case, and then we study its nonstationary version in order to avoid the crucial and not easy choice of the regularization parameter. Note that the iterated version of the weighted-I Tikhonov was already studied in [1].

6.1. Stationary iterated weighted-II Tikhonov. We start by stating a general definition of stationary iterated Tikhonov methods with operator-dependent norms.

DEFINITION 6.1. *Let* $f : [0, \sigma_1] \to [0, \infty)$ *be a Borel-measurable function such that* $F_\alpha(\sigma) = \frac{\sigma^2}{\sigma^2 + \alpha f(\sigma^2)}$ *satisfies conditions* (2.2a)*,* (2.2b)*, and* (2.2c)*. We define the* stationary iterated general weighted Tikhonov method *as*

$$(6.1) \qquad \begin{cases} x^0_{\alpha_0, f} := 0, \\ [K^*K + \alpha f(K^*K)] x^n_{\alpha, f} := K^*y + \alpha f(K^*K) x^{n-1}_{\alpha, f}, \end{cases}$$

or equivalently,

$$\begin{cases} x^0_{\alpha_0, f} := 0, \\ x^n_{\alpha, f} := \operatorname{argmin}_{x \in \mathcal{X}} \left\{ \|Kx - y\|^2 + \alpha \|x - x^{n-1}_{\alpha, f}\|^2_{\sqrt{f(K^*K)}} \right\}. \end{cases}$$

We define $x^{n,\delta}_{\alpha, f}$ *as the* n*-th iteration of* (6.4) *whenever* y *is replaced by* y^δ*.*

Then, by applying the above definition to the special case $f(\sigma^2) = \left(1 - \left(\frac{\sigma}{\sigma_1} \right)^2 \right)^j$, we have the following new definition.

DEFINITION 6.2 (SIWT-II). *We define the* stationary iterated weighted-II Tikhonov method (SIWT-II) *as a stationary iterated general weighted Tikhonov method with the function* $f(\sigma^2) = \left(1 - \left(\frac{\sigma}{\sigma_1} \right)^2 \right)^j$*, namely,*

$$(6.2) \qquad \begin{cases} x^0_{\alpha, j} := 0; \\ \left(K^*K + \alpha \left[I - \frac{K^*K}{\|K^*K\|} \right]^j \right) x^n_{\alpha, j} := K^*y + \alpha \left[I - \frac{K^*K}{\|K^*K\|} \right]^j x^{n-1}_{\alpha, j}, \end{cases}$$

with $\alpha > 0$ *and* $j \in \mathbb{N}$*, or equivalently,*

$$\begin{cases} x^0_{\alpha, j} := 0 \\ x^n_{\alpha, j} := \operatorname{argmin}_{x \in \mathcal{X}} \left\{ \|Kx - y\|^2 + \alpha \|x - x^{n-1}_{\alpha, j}\|^2_{\sqrt{(f(K^*K))}} \right\}. \end{cases}$$

We define $x^{n,\delta}_{\alpha, j}$ *as the* n*-th iteration of* (6.2) *whenever* y *is replaced by* y^δ*.*

PROPOSITION 6.3. *For any given* $n, j \in \mathbb{N}$*, the SIWT-II in* (6.2) *is a filter-based regularization method with filter function*

$$F^{(n)}_{\alpha, j}(\sigma) = \frac{\left(\sigma^2 + \alpha \left[1 - \left(\frac{\sigma}{\sigma_1} \right)^2 \right]^j \right)^n - \left(\alpha \left[1 - \left(\frac{\sigma}{\sigma_1} \right)^2 \right]^j \right)^n}{\left(\sigma^2 + \alpha \left[1 - \left(\frac{\sigma}{\sigma_1} \right)^2 \right]^j \right)^n} = 1 - (1 - F_{\alpha, j}(\sigma))^n.$$

Moreover, the method is of optimal order under the a-priori assumption $x^\dagger \in \mathcal{X}_{\nu, \rho}$*, for* $j \in \mathbb{N}$ *and* $0 < \nu \le 2n$*, with the best convergence rate* $\|x^\dagger - x^{n,\delta}_{\alpha, j}\| = O(\delta^{\frac{2n}{1+2n}})$*, which is obtained for* $\alpha = \left(\frac{\delta}{\rho} \right)^{\frac{2n}{1+\nu}}$ *with* $\nu = 2n$*.*

REMARK 6.4. Let us observe that for $n = 1$ we recover the convergence rate of $O\left(\delta^{\frac{2}{3}}\right)$ in Proposition 3.6.

Proof. Multiplying both sides of (6.1) by $(K^*K + \alpha f(K^*K))^{n-1}$ and iterating the process and since $f(K^*K)$ and K^*K commute, see [23, Section 12.24, p. 326], we get

$$(K^*K + \alpha f(K^*K))^n x_{\alpha,f}^n = \left\{ \sum_{k=0}^{n-1} [\alpha f(K^*K)]^k (K^*K + \alpha f(K^*K))^{n-1-k} \right\} K^* y$$

$$= [(K^*K + \alpha f(K^*K))^n - (\alpha f(K^*K))^n](K^*K)^{-1} K^* y,$$

where we have used the well-known formula $(B^n - A^n) = (B - A)\left(\sum_{k=1}^{n-1} A^k B^{n-1-k}\right)$. Therefore, the filter function in (2.1) is equal to

$$F_{\alpha,f}^{(n)}(\sigma) = \frac{\left(\sigma^2 + \alpha f(\sigma^2)\right)^n - \left(\alpha f(\sigma^2)\right)^n}{\left(\sigma^2 + \alpha f(\sigma^2)\right)^n}.$$

Setting

$$f(\sigma^2) = \left[1 - \left(\frac{\sigma}{\sigma_1}\right)^2\right]^j,$$

we obtain the filter function for the SIWT-II, i.e.,

$$F_{\alpha,j}^{(n)}(\sigma) = \frac{\left(\sigma^2 + \alpha\left[1 - \left(\frac{\sigma}{\sigma_1}\right)^2\right]^j\right)^n - \left(\alpha\left[1 - \left(\frac{\sigma}{\sigma_1}\right)^2\right]^j\right)^n}{\left(\sigma^2 + \alpha\left[1 - \left(\frac{\sigma}{\sigma_1}\right)^2\right]^j\right)^n} = 1 - (1 - F_{\alpha,j}(\sigma))^n.$$

Again, condition (2.2c) is straightforward to verify. Moreover, we recover the following relation

$$F_{\alpha,f}^{(n)}(\sigma) = \frac{\left(\sigma^2 + \alpha f(\sigma^2)\right)^n - \left(\alpha f(\sigma^2)\right)^n}{\left(\sigma^2 + \alpha f(\sigma^2)\right)^n} = 1 - (1 - F_{\alpha,f}(\sigma))^n$$

$$= F_{\alpha,f}(\sigma) \frac{1 - (1 - F_{\alpha,f}(\sigma))^n}{1 - (1 - F_{\alpha,f}(\sigma))} = F_{\alpha,f}(\sigma) \sum_{k=0}^{n-1} (1 - F_{\alpha,f}(\sigma))^k,$$

from which it follows that

$$F_{\alpha,f}(\sigma) \leq F_{\alpha,f}^{(n)}(\sigma) \leq n F_{\alpha,f}(\sigma).$$

Therefore, when we specialize the general function f in the above inequalities to that in Definition 3.2, conditions (2.2a), (2.2b), and (2.6a) follow immediately for every $j \in \mathbb{N}$ thanks to Proposition 3.6. Finally, condition (2.6b) becomes

$$\sup_{\sigma \in [0,\sigma_1]} \left| \left(\frac{\alpha\left[1 - \left(\frac{\sigma}{\sigma_1}\right)^2\right]^j}{\sigma^2 + \alpha\left[1 - \left(\frac{\sigma}{\sigma_1}\right)^2\right]^j} \right)^n \sigma^\nu \right| \leq \sup_{\sigma \in [0,\sigma_1]} \left| \left(\frac{\alpha}{\sigma^2 + \alpha\left[1 - \left(\frac{\sigma}{\sigma_1}\right)^2\right]^j} \right)^n \sigma^\nu \right|$$

$$= \sup_{\sigma \in [0,\sigma_1]} \left| \left(\frac{\alpha \sigma^{\nu/n}}{\sigma^2 + \alpha\left[1 - \left(\frac{\sigma}{\sigma_1}\right)^2\right]^j} \right)^n \right|,$$

and then by using the same approach as in (3.7), it is easy to verify that the last term in the above inequality is bounded by $\alpha^{\beta\nu}$ with $\beta = 1/2$ if and only if $0 < \nu \le 2n$. The remaining assertions follow from an application of Theorem 2.5. $\quad\Box$

6.2. Nonstationary iterated weighted Tikhonov regularization. We introduce a nonstationary version of the iteration (6.2). We study convergence, and we prove that the new iteration is a regularization method.

DEFINITION 6.5. *Let $\{\alpha_n\}_{n\in\mathbb{N}} \subset \mathbb{R}_{>0}$ be a sequence of positive real numbers, and let $f_n : [0, \sigma_1] \to [0, \infty)$ be a sequence of Borel-measurable functions such that $F_{\alpha,n}(\sigma) = \frac{\sigma^2}{\sigma^2 + \alpha f_n(\sigma^2)}$ satisfies (2.2a), (2.2b), and (2.2c) for every n. We introduce the following* nonstationary generalized weighted Tikhonov method,

$$(6.3) \qquad \begin{cases} x^0_{\alpha_0, f_0} := 0, \\ [K^*K + \alpha_n f_n(K^*K)]\, x^n_{\alpha_n, f_n} := K^*y + \alpha_n f_n(K^*K)\, x^{n-1}_{\alpha_{n-1}, f_{n-1}}, \end{cases}$$

or equivalently,

$$\begin{cases} x^0_{\alpha_0, f_0} := 0, \\ x^n_{\alpha_n, f_n} := \operatorname{argmin}_{x\in\mathcal{X}} \left\{ \|Kx - y\|^2 + \alpha_n \|x - x^{n-1}_{\alpha_{n-1}, f_{n-1}}\|^2_{\sqrt{f_n(K^*K)}} \right\}. \end{cases}$$

DEFINITION 6.6 (NSWIT-II). *Let $\{\alpha_n\}_{n\in\mathbb{N}} \subset \mathbb{R}_{>0}$ and $\{j_n\}_{n\in\mathbb{N}} \subset \mathbb{N}$ be sequences of positive real numbers and integers, respectively. We define the* nonstationary iterated weighted-II Tikhonov method (NSIWT-II) *as the nonstationary generalized weighted Tikhonov method with $f_n(\sigma^2) = \left(1 - \frac{\sigma^2}{\sigma_1^2}\right)^{j_n}$, namely,*

$$(6.4) \qquad \begin{cases} x^0_{\alpha_0, j_0} := 0, \\ \left[K^*K + \alpha_n\left(I - \dfrac{K^*K}{\|K^*K\|}\right)^{j_n}\right] x^n_{\alpha_n, j_n} \\ \qquad\qquad := K^*y + \alpha_n\left(I - \dfrac{K^*K}{\|K^*K\|}\right)^{j_n} x^{n-1}_{\alpha_{n-1}, j_{n-1}}, \end{cases}$$

or equivalently,

$$\begin{cases} x^0_{\alpha_0, j_0} := 0, \\ x^n_{\alpha_n, j_n} := \operatorname{argmin}_{x\in\mathcal{X}} \left\{ \|Kx - y\|^2 + \alpha_n \|x - x^{n-1}_{\alpha_{n-1}, j_{n-1}}\|^2_{\sqrt{f_n(K^*K)}} \right\}. \end{cases}$$

6.2.1. Convergence analysis. We are interested in conditions for the sequence $\{\alpha_n\}$ such that the iteration (6.4) converges.

Let $\{E_{\sigma^2}\}_{\sigma^2 \in \sigma(K^*K)}$ be the spectral decomposition of the self-adjoint operator K^*K, and let $\sigma(K^*K)$ denote its spectrum. Then from well-known facts from functional analysis [23], 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 Borel-measurable function and $\langle Ex_1, x_2\rangle$ is a regular complex Borel measure for every $x_1, x_2 \in \mathcal{X}$. Hereafter, without loss of generality, we assume $\sigma_1 = 1$, i.e., $\|K\| = 1$.

THEOREM 6.7. *For every $x^\dagger \in \mathcal{X}$, the method (6.3) converges to $x^\dagger$ as $n \to \infty$ if and only if*

$$\sum_{k=1}^n \frac{\sigma^2}{\sigma^2 + \alpha_k f_k(\sigma^2)}$$

diverges for every $\sigma \in \sigma(K) \setminus \{0\}$.

Proof. Rewriting equation (6.3) using $y = Kx^\dagger$, we have

$$
\begin{aligned}
x_{\alpha_n, f_n}^n &= \left[K^*K + \alpha_n f_n \left(K^*K \right) \right]^{-1} K^* K x^\dagger \\
&\qquad + \alpha_n \left[K^*K + \alpha_n f_n \left(K^*K \right) \right]^{-1} f_n \left(K^*K \right) x_{\alpha_{n-1}, f_{n-1}}^{n-1} \\
&= \left\{ I - \alpha_n \left[K^*K + \alpha_n f_n \left(K^*K \right) \right]^{-1} f_n \left(K^*K \right) \right\} x^\dagger \\
&\qquad + \alpha_n \left[K^*K + \alpha_n f_n \left(K^*K \right) \right]^{-1} f_n \left(K^*K \right) x_{\alpha_{n-1}, f_{n-1}}^{n-1},
\end{aligned}
$$

from which it follows that

$$
\begin{aligned}
x^\dagger - x_{\alpha_n, f_n}^n &= \alpha_n \left[K^*K + \alpha_n f_n \left(K^*K \right) \right]^{-1} f_n \left(K^*K \right) \left(x^\dagger - x_{\alpha_{n-1}, f_{n-1}}^{n-1} \right) \\
&= \prod_{k=1}^n \alpha_k \left[K^*K + \alpha_k f_k \left(K^*K \right) \right]^{-1} f_k \left(K^*K \right) x^\dagger,
\end{aligned}
\tag{6.5}
$$

where for convenience we set $x_{\alpha_0, f_0}^0 := 0$. As a consequence, the method converges for every $x^\dagger$ if and only if

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

for every $x^\dagger \in \mathcal{X}$, i.e., if and only if

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

for every Borel-measure $\langle E x^\dagger, x^\dagger \rangle$ induced by $x^\dagger \in \mathcal{X}$. Since

$$
\left| \prod_{k=1}^n \frac{\alpha_k f_k \left(\sigma^2 \right)}{\sigma^2 + \alpha_k f_k \left(\sigma^2 \right)} \right|^2 \le 1
$$

for every n and since

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

the theorem on dominated convergence implies the following equality

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

Hence, the method is convergent for every $x^\dagger \in \mathcal{X}$ if and only if

$$
\prod_{k=1}^\infty \frac{\alpha_k f_k \left(\sigma^2 \right)}{\sigma^2 + \alpha_k f_k \left(\sigma^2 \right)} = \prod_{k=1}^\infty \left(1 - \frac{\sigma^2}{\sigma^2 + \alpha_k f_k \left(\sigma^2 \right)} \right) = 0,
$$

for $\langle Ex^{\dagger}, x^{\dagger}\rangle$-a.e. σ^2 and every induced Borel measure $\langle Ex^{\dagger}, x^{\dagger}\rangle$, i.e., for every $\sigma \in \sigma(K)\backslash\{0\}$. Now using the fact that

$$\prod_{k=1}^{\infty} (1 - t_k) = 0 \qquad \text{if and only if} \qquad \sum_{k=1}^{\infty} t_k = \infty,$$

whenever $\{t_k\}_{k\in\mathbb{N}}$ is a sequence of positive real numbers such that $0 \le t_k < 1$ ([22]), the assertion follows.   □

COROLLARY 6.8. *For every $x^{\dagger} \in \mathcal{X}$, the method (6.4) converges to $x^{\dagger}$ as $n \to \infty$ if and only if*

$$\sum_{k=1}^{n} \frac{\sigma^2}{\sigma^2 + \alpha_k \left[1 - \left(\frac{\sigma}{\sigma_1}\right)^2\right]^{j_k}}$$

diverges for every $\sigma \in \sigma(K) \setminus \{0\}$.

COROLLARY 6.9. *If $\sum_{k=1}^{n} \alpha_k^{-1}$ diverges, then the NSIWT-II method converges.*

Proof. Let us observe that

$$\sum_{k=1}^{\infty} \frac{\sigma^2}{\sigma^2 + \alpha_k} \le \sum_{k=1}^{\infty} \frac{\sigma^2}{\sigma^2 + \alpha_k \left[1 - \left(\frac{\sigma}{\sigma_1}\right)^2\right]^{j_k}}.$$

Then, if the left-hand side of the above inequality diverges for every $\sigma \in \sigma(K) \setminus \{0\}$, then the result follows.

If $\sum_{k=1}^{\infty} \alpha_k^{-1} = \infty$, we can possibly have three different cases:

$$\lim_{k\to\infty} \alpha_k \in [0, \infty), \quad \nexists \lim_{k\to\infty} \alpha_k, \quad \text{or} \quad \lim_{k\to\infty} \alpha_k = \infty.$$

In the first two cases, $\frac{\sigma^2}{\sigma^2 + \alpha_k} \nrightarrow 0$ for every $\sigma > 0$, and then the corresponding series diverges. In the latter case instead, $\alpha_k^{-1} \sim c_\sigma \frac{\sigma^2}{\sigma^2 + \alpha_k}$ for every $\sigma > 0$, and hence, the series $\sum_{k=1}^{n} \alpha_k^{-1}$ and $\sum_{k=1}^{n} \frac{\sigma^2}{\sigma^2 + \alpha_k}$ converge or diverge simultaneously by the asymptotic comparison test. Then, by $\sum_{k=1}^{\infty} \alpha_k^{-1} = \infty$, we deduce that $\sum_{k=1}^{\infty} \frac{\sigma^2}{\sigma^2 + \alpha_k}$ diverges for every $\sigma > 0$, and the NSIWT-II method converges.   □

Now, we investigate the convergence rate of NSIWT-II.

THEOREM 6.10. *Let $\{x^n_{\alpha_n, f_n}\}_{n\in\mathbb{N}}$ be a convergent sequence of the nonstationary generalized weighted Tikhonov method (6.3) 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*

$$(6.6a) \qquad \lim_{n\to\infty} \vartheta_n \sigma^\nu \prod_{k=1}^{n} \left(1 - \frac{\sigma^2}{\sigma^2 + \alpha_k f_k(\sigma^2)}\right) = 0 \qquad \text{for every } \sigma \in \sigma(K) \setminus \{0\},$$

(6.6b)

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

then

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

Proof. From equation (6.5), for $x^\dagger \in \mathcal{X}_\nu$, we have

$$\lim_{n\to\infty} \vartheta_n \|x^\dagger - x^n_{\alpha_n, f_n}\|$$

$$= \lim_{n\to\infty} \left[\int_{\sigma(K^*K)} \left| \vartheta_n \sigma^\nu \prod_{k=1}^n \left(1 - \frac{\sigma^2}{\sigma^2 + \alpha_k f_k(\sigma^2)} \right) \right|^2 d\langle E_{\sigma^2}\omega, \omega \rangle \right]^{1/2}$$

$$= \left[\int_{\sigma(K^*K)} \left| \lim_{n\to\infty} \vartheta_n \sigma^\nu \prod_{k=1}^n \left(1 - \frac{\sigma^2}{\sigma^2 + \alpha_k f_k(\sigma^2)} \right) \right|^2 d\langle E_{\sigma^2}\omega, \omega \rangle \right]^{1/2},$$

by (6.6b) and the dominated convergence theorem. Now the assertion follows from hypothesis (6.6a). $\quad\square$

For example, in the case of nonstationary iterated Tikhonov, i.e., for $f_k(\sigma^2) \equiv 1$, setting $\beta_n = \sum_{k=1}^n (1 + \alpha_k)^{-1}$, the conditions (6.6a) and (6.6b) are satisfied with $\vartheta_n = \beta_n^{\nu/2}$. Anyway, this is not always the best convergence rate. Indeed, in practice a common choice for $\{\alpha_k\}$ is the geometric sequence $\alpha_k = \alpha_0 q^k$ with $0 < q < 1$ and α_0 fixed, which provides a convergence rate of the form $O(q^n)$. We invite the interested reader to study more details in the following papers: [2], [11] and [1, Corollary 27].

Contextualizing the above theorem to the case of NSIWT-II, we have the following corollary.

COROLLARY 6.11. *If conditions* (6.6a) *and* (6.6b) *are satisfied for*

$$f_k(\sigma^2) = \left[1 - \left(\frac{\sigma}{\sigma_1} \right)^2 \right]^{j_k},$$

then we have

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

6.2.2. Analysis of convergence for perturbed data. Now, we consider the convergence of the NSIWT-I/II methods when the initial datum y is perturbed. Again, we initially prove a more general statement involving the method (6.3) with initial datum y^δ, and then the convergence results for perturbed data in the NSIWT-I/II cases will follow as a corollary. We use the notation $x^{n,\delta}_{\alpha_n, f_n}$ for the solution of the method (6.3) with perturbed initial datum y^δ.

THEOREM 6.12. *Under the assumptions of Theorem 6.7, let $f_k : [0, \sigma_1] \to \mathbb{R}$ be such that*

$$(6.7) \qquad\qquad \sup_{\sigma \in [0,1]} \left| \frac{\alpha_k}{\sigma^2 + \alpha_k f_k(\sigma^2)} \right| \leq \gamma_k.$$

If $\{\delta_n\}$ is a sequence convergent to 0 with $\delta_n \geq 0$ such that

$$(6.8) \qquad\qquad \lim_{n\to\infty} \delta_n \cdot \sum_{k=1}^n \alpha_k^{-1}\gamma_k = 0,$$

then,

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

Proof. From the definition of the method (6.3), for every given $m \leq n$, we find that

$$
\begin{aligned}
x^{m,\delta_n}_{\alpha_m,f_m} &= [K^*K + \alpha_m f_m (K^*K)]^{-1} \left(K^* y^{\delta_n} + \alpha_m f_m (K^*K) x^{m-1,\delta_n}_{\alpha_{m-1},f_{m-1}} \right) \\
&= \left\{ I - \alpha_m [K^*K + \alpha_m f_m (K^*K)]^{-1} f_m (K^*K) \right\} x^\dagger \\
&\quad + \alpha_m [K^*K + \alpha_m f_m (K^*K)]^{-1} f_m (K^*K) x^{m-1,\delta_n}_{\alpha_{m-1},f_{m-1}} \\
&\quad + [K^*K + \alpha_m f_m (K^*K)]^{-1} K^*(y^{\delta_n} - y),
\end{aligned}
$$

thus,

$$
\begin{aligned}
x^\dagger - x^{m,\delta_n}_{\alpha_m,f_m} &= \alpha_m [K^*K + \alpha_m f_m (K^*K)]^{-1} f_m (K^*K) (x^\dagger - x^{m-1,\delta_n}_{\alpha_{m-1},f_{m-1}}) \\
&\quad - [K^*K + \alpha_m f_m (K^*K)]^{-1} K^*(y^{\delta_n} - y).
\end{aligned}
$$

Hence, by induction, setting $g_i(K^*K) = \alpha_i [K^*K + \alpha_i f_i (K^*K)]^{-1} f_i(K^*K)$ for every fixed n, we have

$$
\begin{aligned}
&x^\dagger - x^{n,\delta_n}_{\alpha_n,f_n} \\
&= \prod_{k=1}^n g_k(K^*K)x^\dagger - \sum_{k=1}^n \left(\prod_{i=k+1}^n g_i(K^*K) \right) [K^*K + \alpha_k f_k (K^*K)]^{-1} K^*(y^{\delta_n} - y).
\end{aligned}
$$

Since

$$
\|g_i(K^*K)\| = \sup_{\sigma \in [0,1]} \left| \frac{\alpha_i f_i(K^*K)}{\sigma^2 + \alpha_i f_i(K^*K)} \right| \leq 1, \qquad \|K^*\| = 1,
$$

and

$$
\|\alpha_k [K^*K + \alpha_k f_k (K^*K)]^{-1}\| = \sup_{\sigma \in [0,1]} \left| \frac{\alpha_k}{\sigma^2 + \alpha_k f_k (\sigma^2)} \right| \leq \gamma_k,
$$

it follows that

$$
\begin{aligned}
&\|x^\dagger - x^{n,\delta_n}_{\alpha_n,f_n}\| \\
&\leq \| \prod_{k=1}^n g_k(K^*K)x^\dagger \| \\
&\quad + \sum_{k=1}^n \alpha_k^{-1} \prod_{i=k+1}^n \|g_i(K^*K)\| \|\alpha_k [K^*K + \alpha_k f_k (K^*K)]^{-1}\| \|K^*\| \|y^{\delta_n} - y\| \\
&\leq \|x^\dagger - x^n_{\alpha_n,f_n}\| + \delta_n \sum_{k=1}^n \alpha_k^{-1} \gamma_k.
\end{aligned}
$$

Finally, by Theorem 6.7, $\|x^\dagger - x^n_{\alpha_n,f_n}\| \to 0$ as $n \to \infty$, and therefore by equation (6.8), $\|x^\dagger - x^{n,\delta_n}_{\alpha_n,f_n}\| \to 0$ for $n \to \infty$. $\quad\square$

COROLLARY 6.13. *Under the assumptions of Corollary 6.9, let $\{\alpha_k\}, \{j_k\}$ be such that $j_k \leq \alpha_k^{-1}$. If $\{\delta_n\}$ is a sequence convergent to 0 with $\delta_n \geq 0$ such that*

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

then the NSIWT-II method with perturbed data is convergent, that is,

$$\lim_{n \to \infty} \|x^{\dagger} - x_{\alpha_n, j_n}^{n, \delta_n}\| = 0.$$

Proof. Applying Theorem 6.12 with

$$f_k(\sigma^2) = \left(1 - \sigma^2\right)^{j_k},$$

we have that

$$\sup_{\sigma \in [0,1]} \left| \frac{\alpha_k}{\sigma^2 + \alpha_k f_k(\sigma^2)} \right| = \sup_{\sigma \in [0,1]} \left| \frac{\alpha_k}{\sigma^2 + \alpha_k \left(1 - \sigma^2\right)^{j_k}} \right|.$$

Setting $\frac{\alpha_k}{\sigma^2 + \alpha_k (1 - \sigma^2)^{j_k}} = h_k(\sigma)$, from $j_k \leq \alpha_k^{-1}$ it follows that

$$h_k'(\sigma) = -\frac{2\alpha_k \sigma \left[1 - \alpha_k j_k \left(1 - \sigma^2\right)^{j_k - 1}\right]}{\left[\sigma^2 + \alpha_k \left(1 - \sigma^2\right)^{j_k}\right]^2} \leq 0 \qquad \text{for every } k \in \mathbb{N}, \sigma \in [0, 1].$$

Therefore

$$\sup_{\sigma \in [0,1]} \left| \frac{\alpha_k}{\sigma^2 + \alpha_k \left(1 - \sigma^2\right)^{j_k}} \right| \leq 1,$$

and the hypothesis (6.7) is satisfied and the assertion follows at once. □

6.3. Iterated mixed method. With the same idea in mind that led us to introduce the filter method in Definition 3.4, we give the following definition of an iterated method which combines both the iterated weighted-I and the iterated weighted-II methods.

DEFINITION 6.14 (NSM). *Let $\{\alpha_n\}_{n \in \mathbb{N}}, \{r_n\}_{n \in \mathbb{N}} \subset \mathbb{R}_{>0}$, and $\{j_n\}_{n \in \mathbb{N}} \subset \mathbb{N}$ be sequences of positive real numbers and integers, respectively. We define the* nonstationary iterated mixed method (NSM) *as a nonstationary generalized Tikhonov method (see Definition 6.5) with*

$$f_n(\sigma^2) = \left(1 - \frac{\sigma^2}{\sigma_1^2}\right)^{j_n} \sigma^{1 - r_n}.$$

We skip all the convergence analyses as they can be studied and recovered without great efforts by adapting the proofs in Proposition 6.3, Corollaries 6.8, 6.9, 6.13, and Theorem 6.10.

7. Numerical examples. Here we present some numerical examples, in the direct and iterated cases, spanning all the methods we discussed in the previous sections.

To produce our results we have used Matlab R2015a on a desktop PC provided with an Intel iCore i5-4460 processor with 8 GB of RAM running Windows 8.1. We add the "noise-vector" η to the noise-free right-hand side vector y, where η has normally distributed
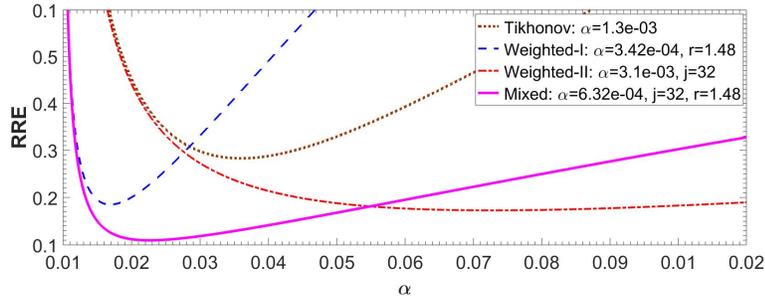
FIGURE 7.1. *RRE vs α values for the Heat problem with $\kappa = 3$.*

pseudorandom entries with mean zero in all examples and is normalized to correspond to a chosen noise-level

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

As a stopping criterion for the iterative methods, we use the discrepancy principle [12] that terminates the iterative method at the iteration

$$\hat{k} = \min_{k} \left\{ k : \ \|y^{\delta} - Kx_k\| \leq \tau\xi \right\}, \qquad \tau = 1.01,$$

i.e., this criterion stops the iterations when the norm of the residual reaches the norm of the noise so that the latter does not dominate the reconstructed solution. All the iterative methods are initialized with the zero vector.

To compare the restorations with the different methods, we consider both the visual representation and the relative restoration error (RRE), that is,

$$\frac{\|x^{\dagger} - \hat{x}\|}{\|x^{\dagger}\|},$$

for the computed approximations $\hat{x}$.

7.1. Example 1. The following test case is *Heat* from the toolbox Regularization Tools by Hansen [13] with the parameter $\kappa = 3$ and using 300 points. We add a noise vector with $\xi = 0.05$. The comparisons between direct methods have been made by choosing the best set of parameters for each method.
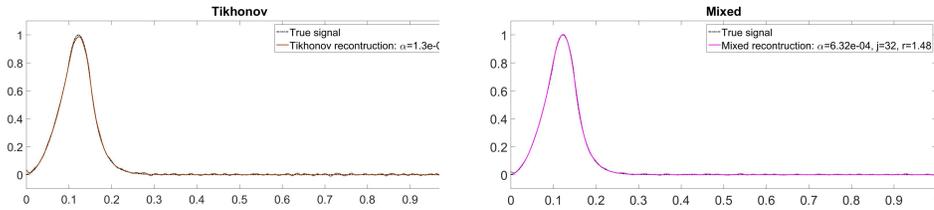
In Table 7.1 we can observe that the mixed method gives the lowest relative restoration error.

TABLE 7.1
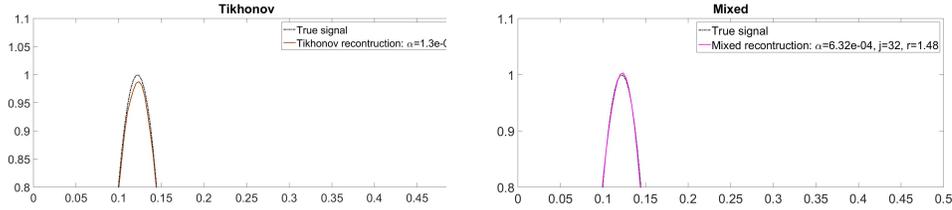*Heat problem. Parameters and RRE for the direct methods.*

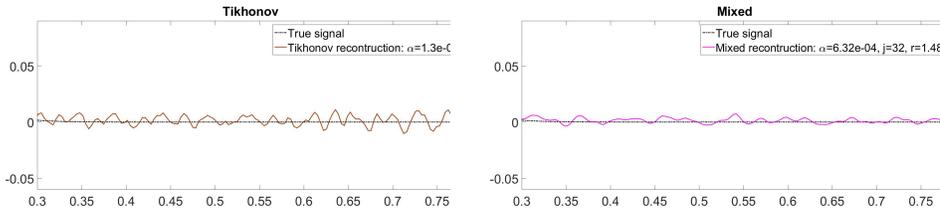| **Method**	α	j	r	**RRE**
Tikhonov	$1.3\mathrm{e} - 03$	0	1	0.0241
Weighted-I	$3.42\mathrm{e} - 04$	0	1.48	0.0193
Weighted-II	$3.1\mathrm{e} - 03$	32	1	0.0186
Mixed	$6.32\mathrm{e} - 04$	32	1.48	**0.0155**
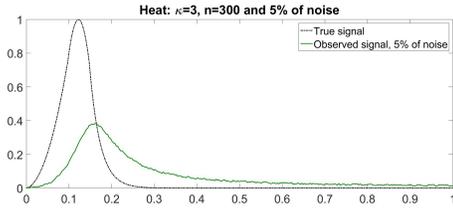
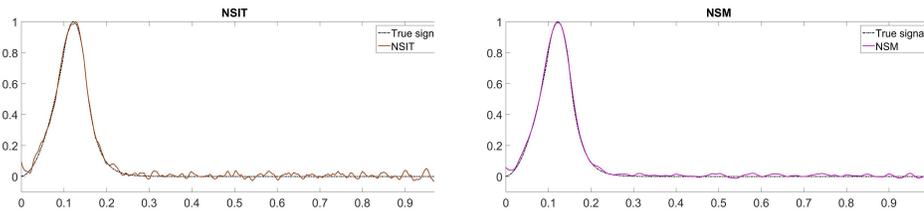*(a) True signal and observed signal corrupted by 5% of noise.*

(b) Graphical comparison between Tikhonov and mixed methods.

(c) Detail: the head of the signal.

(d) Detail: the tail of the signal.

(e) Graphical comparison between NSIT and NSM.

FIGURE 7.2. *Heat problem. Comparison among some regularization methods, computed solutions and details.*

Since the true signal $x^\dagger$ is a smooth function, the best parameter r is larger than 1. This is consistent with the theory: indeed according to Proposition 4.1, for $r \geq 1$ the Weighted-I method oversmooths the approximate solution more when compared to the standard Tikhonov method. In Figure 7.1 we plot the α-values against the RRE for every method. We can detect an empirical behavior for the Weighted-II method which was already observed in [17], i.e., its

curve stays below the Tikhonov curve as well as below all the other ones. Indeed, the slope of the curve for the Weighted-II method in a right neighborhood of its minimum is less than the slopes of the curves for all the other methods calculated in a right neighborhood centered at their minimums. This means that an overestimation of the best parameter α would affect the RRE less.

Interestingly, it appears that the mixed method combines the best features of both the Weighted-I and the Weighted-II, or more precisely, the Weighted-II method enhances the Weighted-I method. This is clearly seen in both Figures 7.1 and 7.2. Figures 7.2(c)–(d) highlight the differences between the standard Tikhonov method and the mixed method: the *head* of the signal is reconstructed better by the mixed method, and at the same time it dampens the rough oscillations that appear in the *tail* of the reconstructed signal by the Tikhonov method.

Finally, we propose a comparison between the nonstationary iterative versions of the preceding methods. We denote by NSIT the nonstationary iterated Tikhonov method and with NSM the nonstationary iterated mixed method. In Table 7.2 we report the RRE values and the corresponding sequences of parameters that we use for the methods. Even in this case, the NSM method produces a better approximation, and in Figure 7.2(e) a comparison between the approximate solutions derived from NSIT and from NSM is reported. We observe that the behavior of the nonstationary iterative methods is not significantly affected by the choice of α_0 and q. In practice, in the iterative cases we did not have to make any particular choice on the parameters. Of course, a lower RRE would be obtained by stopping the iterative method at the best iteration corresponding to the minimum RRE. Nevertheless, despite a larger RRE, by visual inspection of Figure 7.2, the peak is restored more accurately with NSM than with the Tikhonov method.

TABLE 7.2
Heat problem. RRE values for the nonstationary iterative methods and iteration numbers between brackets.

Method	$\alpha_n = \alpha_0 q^n$	j_n	r_n	RRE
NSIT	$\alpha_0 = 0.01, q = 0.7$	0	1	0.0733(4)
NSWIT-I	$\alpha_0 = 0.01, q = 0.7$	0	$1.2 + n/100$	0.0590(3)
NSWIT-II	$\alpha_0 = 0.01, q = 0.7$	$4n$	1	0.0530(3)
NSM	$\alpha_0 = 0.01, q = 0.7$	$4n$	$1.2 + n/100$	**0.0485**(3)

7.2. Example 2. For completeness and in order to give an overall view, here we report a table that summarizes different test problems from the toolbox Regularization Tools by Hansen [13]. We have studied two different cases, where for each test we add a noise vector with $\xi = 0.05$ and $\xi = 0.01$, respectively. We compare only the direct methods, and the comparisons have been made by choosing the best set of parameters for each method. In most of the cases the mixed method provides the best RRE, and when it does not, it improves the reconstructed solution compared to the cases that use the weighted-I or the weighted-II method alone.

7.3. Example 3. The following test case is *Foxgood* from the toolbox Regularization Tool by Hansen [13] with $n = 300$ points. We added to the observed data a noise vector with $\xi = 0.01$.

In this example we are going to examine only the nonstationary iterative algorithms. As in the previous example, we choose a geometric sequence for the α_n in every iterated method, while the j_n and the r_n are monotonically increasing sequences. The numerical results in

TABLE 7.3
RRE values for different problems from the toolbox Regularization Tools. Here $n = 300$.

Problem	Method	RRE	
		$\xi = 5\%$	$\xi = 1\%$
baart(n)	Tikhonov: $\alpha = 2.2\mathrm{e} - 05$	0.5560	0.1174
	Weighted-I: $\alpha = 1.2\mathrm{e} - 05, r = 0.85$	0.4672	0.0870
	Weighted-II: $\alpha = 2\mathrm{e} - 06, j = 32$	0.3390	0.0874
	Mixed: $\alpha = 2\mathrm{e} - 06, r = 0.85, j = 32$	**0.1804**	**0.0757**
deriv2$(n, 3)$	Tikhonov: $\alpha = 2.2\mathrm{e} - 05$	0.1555	0.0491
	Weighted-I: $\alpha = 6.2\mathrm{e} - 05, r = 0.8$	0.1425	0.0449
	Weighted-II: $\alpha = 3.2\mathrm{e} - 05, j = 16$	**0.1187**	0.0477
	Mixed: $\alpha = 7.2\mathrm{e} - 05, r = 0.8, j = 16$	0.1327	**0.0430**
gravity$(n, 3)$	Tikhonov: $\alpha = 5.6\mathrm{e} - 02$	0.1166	0.0858
	Weighted-I: $\alpha = 6.8\mathrm{e} - 02, r = 0.7$	0.0968	**0.0738**
	Weighted-II: $\alpha = 1.7\mathrm{e} - 02, j = 32$	0.1145	0.0779
	Mixed: $\alpha = 4.6\mathrm{e} - 02, r = 0.7, j = 32$	**0.0945**	0.0752
phillips(n)	Tikhonov: $\alpha = 2.44\mathrm{e} - 02$	0.0645	0.0244
	Weighted-I: $\alpha = 1.18\mathrm{e} - 02, r = 1.2$	0.0628	0.0182
	Weighted-II: $\alpha = 1.35\mathrm{e} - 02, j = 32$	0.0765	0.0177
	Mixed: $\alpha = 7.9\mathrm{e} - 03, r = 1.2, j = 32$	**0.0623**	**0.0144**

(a) *True signal and observed signal corrupted by 1% of noise.*

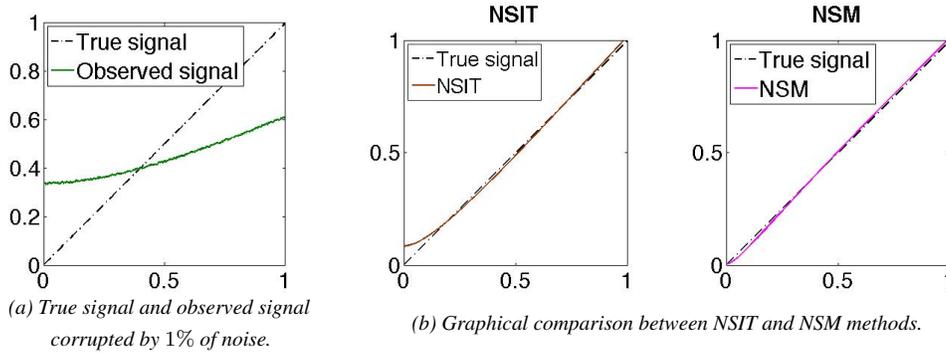(b) *Graphical comparison between NSIT and NSM methods.*

FIGURE 7.3. *Foxgood problem. Computed solutions with NSIT and NSM methods.*

Table 7.4 and the plots in Figure 7.3 show that the NSM method provides again the best approximate solution.

TABLE 7.4
Foxgood problem. RRE values for the nonstationary iterative methods and iteration numbers between brackets.

Method	$\alpha_n = \alpha_0 q^n$	j_n	r_n	RRE
NSIT	$\alpha_0 = 0.05, q = 0.8$	0	1	0.0320(15)
NSWIT-I	$\alpha_0 = 0.05, q = 0.8$	0	$1.2 + n/100$	0.0195(21)
NSWIT-II	$\alpha_0 = 0.05, q = 0.8$	$4n$	1	0.0231(17)
NSM	$\alpha_0 = 0.05, q = 0.8$	$4n$	$1.2 + n/100$	**0.0188**(24)

8. Conclusions. We have studied fundamental properties such as convergence, smoothing effects, and saturation for both direct and iterative methods that arise when using general

operator-dependent seminorms in the penalty term. In particular, we have seen that the Weighted-II method performs well as a switch between different regularization effects, and it can be used in combination with other suitable filters to enhance the reconstruction. Clearly, the combination of more filters requires the estimation of more parameters, and techniques like those proposed in [8] should be considered.

As future work, the filtering operators introduced in this paper could be applied also to nonstationary preconditioned iterative methods like the approximated iterated Tikhonov methods with a general penalty term proposed in [3], which extends the technique introduced in [5] using a generalized Tikhonov method as preconditioner. Moreover, they could be applied as smoothers in multilevel regularization methods like in [6].

REFERENCES

[1] D. BIANCHI, A. BUCCINI, M. DONATELLI, AND S. SERRA-CAPIZZANO, *Iterated fractional Tikhonov regularization*, Inverse Problems, 31 (2015), Art. Id. 055005, 34 pages.

[2] M. BRILL AND E. SCHOCK, *Iterative solution of ill-posed problems—a survey*, in Model Optimization in Exploration Geophysics, A. Vogel, ed., Theory Practice Appl. Geophys. 1, Vieweg, Braunschweig, 1987, pp. 13–37.

[3] A. BUCCINI, *Regularizing preconditioners by non-stationary iterated Tikhonov with general penalty term*, Appl. Numer. Math., 116 (2017), pp. 64–81.

[4] A. BUCCINI, M. DONATELLI, AND L. REICHEL, *Iterated Tikhonov regularization with a general penalty term*, Numer. Linear Algebra Appl., 24 (2017), Art. Id. e2089, 12 pages.

[5] M. DONATELLI AND M. HANKE, *Fast nonstationary preconditioned iterative methods for ill-posed problems, with application to image deblurring*, Inverse Problems, 29 (2013), Art. Id. 095008, 16 pages.

[6] M. DONATELLI AND S. SERRA-CAPIZZANO, *Filter factor analysis of an iterative multilevel regularizing method*, Electron. Trans. Numer. Anal., 29 (2007/08), pp. 163–177.
http://etna.ricam.oeaw.ac.at/vol.29.2007-2008/pp163-177.dir/pp163-177.pdf

[7] H. W. ENGL, M. HANKE, AND A. NEUBAUER, *Regularization of Inverse Problems*, Kluwer, Dordrecht, 1996.

[8] S. GAZZOLA AND L. REICHEL, *A new framework for multi-parameter regularization*, BIT, 56 (2016), pp. 919–949.

[9] D. GERTH, E. KLANN, R. RAMLAU, AND L. REICHEL, *On fractional Tikhonov regularization*, J. Inverse Ill-Posed Probl., 23 (2015), pp. 611–625.

[10] C. W. GROETSCH, *The Theory of Tikhonov Regularization for Fredholm Equations of the First Kind*, Pitman, Boston, 1984.

[11] M. HANKE AND C. W. GROETSCH, *Nonstationary iterated Tikhonov regularization*, J. Optim. Theory Appl., 98 (1998), pp. 37–53.

[12] M. HANKE AND P. C. HANSEN, *Regularization methods for large-scale problems*, Surveys Math. Indust., 3 (1993), pp. 253–315.

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

[14] M. E. HOCHSTENBACH, S. NOSCHESE, AND L. REICHEL, *Fractional regularization matrices for linear discrete ill-posed problems*, J. Engrg. Math., 93 (2015), pp. 113–129.

[15] M. E. HOCHSTENBACH AND L. REICHEL, *Fractional Tikhonov regularization for linear discrete ill-posed problems*, BIT, 51 (2011), pp. 197–215.

[16] G. HUANG, L. REICHEL, AND F. YIN, *On the choice of solution subspace for nonstationary iterated Tikhonov regularization*, Numer. Algorithms, 72 (2016), pp. 1043–1063.

[17] T. K. HUCKLE AND M. SEDLACEK, *Tikhonov-Phillips regularization with operator dependent seminorms*, Numer. Algorithms, 60 (2012), pp. 339–353.

[18] E. KLANN AND R. RAMLAU, *Regularization by fractional filter methods and data smoothing*, Inverse Problems, 24 (2008), Art. Id. 025018, 26 pages.

[19] F. LENTI, F. NUNZIATA, C. ESTATICO, AND M. MIGLIACCIO, *Spatial resolution enhancement of earth observation products using an acceleration technique for iterative methods*, IEEE Geosci. Remote Sensing, 12 (2015), pp. 269–273.

[20] A. K. LOUIS, *Inverse und Schlecht Gestellte Probleme*, B. G. Teubner, Stuttgart, 1989.

[21] S. NOSCHESE AND L. REICHEL, *Lavrentiev-type regularization methods for Hermitian problems*, Calcolo, 52 (2015), pp. 187–205.

[22] W. RUDIN, *Real and Complex Analysis*, 3rd ed., McGraw-Hill, New York, 1987.

[23] ———, *Functional Analysis*, McGraw-Hill, New York, 1991.