Doctoral Program
Computational Mathematics
Numerical Analysis and Symbolic Computation

On Fractional Tikhonov Regularization

Daniel Gerth Esther Klann Ronny Ramlau
Lothar Reichel

DK-Report No. 14-06 04 2014

A–4040 LINZ, ALTENBERGERSTRASSE 69, AUSTRIA

Supported by
Austrian Science Fund (FWF) Upper Austria

DK sponsors:

- **Johannes Kepler University Linz** (JKU)
- **Austrian Science Fund** (FWF)
- **Upper Austria**

On fractional Tikhonov regularization

D. Gerth* E. Klann† R. Ramlau ‡†* L. Reichel§

April 2, 2014

Abstract

It is well known that Tikhonov regularization in standard form may determine approximate solutions that are too smooth, i.e., the approximate solution may lack many details that the desired exact solution might possess. Two different approaches, both referred to as fractional Tikhonov methods have been introduced to remedy this shortcoming. This paper investigates the convergence properties of these methods. We show that both methods are order optimal when the regularization parameter is chosen according to the discrepancy principle. The theory developed suggests situations in which the fractional methods yield approximate solutions of higher quality than Tikhonov regularization in standard form. Computed examples that illustrate the behavior of the methods are presented.

1 Introduction

Let A be a linear compact operator between the Hilbert spaces X and Y, and consider the operator equation

$$Ax = b, \qquad x \in X, \quad b \in Y, \tag{1}$$

which we assume to be consistent. We would like to determine the solution of minimal X-norm, which we denote by $x^\dagger$. It can be computed as $x^\dagger = A^\dagger b$, where $A^\dagger$ is the Moore–Penrose pseudoinverse of A. The computation of $x^\dagger$ is an ill-posed problem, because a small perturbation in b may give rise to an arbitrarily large perturbation in $x^\dagger$, or even make the problem unsolvable. Moreover,

*Johannes Kepler University Linz, Doctoral Program "Computational Mathematics", Altenbergerstr. 69, A-4040 Linz, Austria. E-mail: daniel.gerth@dk-compmath.jku.at.Research supported by the Austrian Science Fund (FWF):W1214-N15, Project DK-08.

†Johannes Kepler University Linz, Industrial Mathematics Institute, Altenbergerstr. 69, A-4040 Linz, Austria. E-mail: esther.klann@jku.at.Research supported by FWF grant T529-N18

‡Radon Institute for Computational and Applied mathematics (RICAM), Austrian Academy of Sciences, Altenbergerstr. 69, A-4040 Linz, Austria. E-mail: ronny.ramlau@jku.at.Research in part supported by the Austrian Science Fund (FWF):W1214-N15, Project DK-08

§Department of Mathematical Sciences, Kent State University, Kent, OH 44242, USA. E-mail: reichel@math.kent.edu.Research supported in part by NSF grant DMS-1115385.

the right-hand side function that is available in applications represents data that is contaminated by noise. Thus, instead of b, the error-contaminated function b^δ is available. We assume that a bound for the error

$$\|b^\delta - b\|_Y \leq \delta \tag{2}$$

is known.

Straightforward solution of (1) with b replaced by b^δ generally does not yield a meaningful approximation of $x^\dagger$ because of severe propagation of the error in b^δ into the computed solution. A common remedy, known as *Tikhonov regularization*, is to replace (1) by a penalized least squares problem of the form

$$\min_{x \in X} J_\mu(x) \tag{3}$$

with

$$J_\mu(x) := \|Ax - b^\delta\|_Y^2 + \mu \|x\|_X^2; \tag{4}$$

see, e.g., [2, 7, 9] for discussions and many details on this solution approach. The parameter $\mu > 0$ is referred to as the regularization parameter and determines how sensitive the minimizer x_μ^δ of J_μ is to the error in b^δ and how close x_μ^δ is to the desired solution $x^\dagger$.

Because the bound (2) is known, we may determine a suitable value $\mu > 0$ by the discrepancy principle, i.e., we choose $\mu > 0$ so that

$$\|Ax_\mu^\delta - b^\delta\|_Y = \tau\delta,$$

where $\tau > 1$ is a user-supplied constant that is independent of δ. We refer to x_μ^δ as an *approximate* or *regularized* solution of (1).

The Tikhonov regularization problem (3)-(4) is said to be in *standard form*, because the penalty term is the square of the X-norm of the computed solution. Determining the minimum of (4) is equivalent to solving the normal equations

$$(A^*A + \mu I)x = A^* b^\delta, \tag{5}$$

where A^* denotes the adjoint of A.

It is well known that Tikhonov regularization in standard form typically determines a regularized solution x_μ^δ that is too smooth, i.e., many details of the desired solution $x^\dagger$ typically are not represented by x_μ^δ. This shortcoming led Klann and Ramlau [6] to introduce the fractional Tikhonov regularization method. Subsequently another approach, also referred to as fractional Tikhonov regularization, was investigated by Hochstenbach and Reichel [4]. The latter approach fits into the framework of generalized Tikhonov regularization introduced by Louis [7, Chapter 4]. Application of the fractional approach in [4, 7] to Lavrentiev regularization is discussed in [5].

The method in [4, 7] can be derived by replacing the Y-norm in the fidelity term in (4) by a weighted seminorm

$$\|y\|_W := \|W^{1/2} y\|_Y$$

with

$$W = (AA^*)^{(\alpha-1)/2} \tag{6}$$

for some parameter $0 \leq \alpha \leq 1$, where W is defined with the aid of the Moore–Penrose pseudoinverse of AA^* when $\alpha < 1$. We obtain the minimization problem

$$\min_{x \in X} \widetilde{J}_\mu(x) \tag{7}$$

with

$$\widetilde{J}_\mu(x) := \|Ax - b^\delta\|_W^2 + \mu\|x\|_X^2. \tag{8}$$

We denote the solution of (7)-(8) by $\widetilde{x}_\mu^\delta$. It can be computed by solving the associated normal equations

$$((A^*A)^{(\alpha+1)/2} + \mu I)x = (A^*A)^{(\alpha-1)/2}A^*b^\delta. \tag{9}$$

Oversmoothing in Tikhonov regularization in standard form (which corresponds to $\alpha = 1$) is caused by the fact that b^δ is multiplied by A^*. Letting $0 < \alpha < 1$ reduces oversmoothing.

Klann and Ramlau [6] propose another approach to reduce oversmoothing. They advocate that an approximation of $x^\dagger$ be computed by solving

$$(A^*A + \mu I)^\alpha x = (A^*A)^{\alpha-1}A^*b^\delta \tag{10}$$

for some $0 < \alpha \leq 1$, where $(A^*A)^{\alpha-1}$ is defined with the Moore–Penrose pseudoinverse when $\alpha < 1$. This leads to an interpolation between standard Tikhonov regularization and the generalized inverse. We denote the solution by $\widehat{x}_\mu^\delta$. Also this method simplifies to Tikhonov regularization in standard form when $\alpha = 1$.

The present paper is organized as follows. Section 2 introduces necessary notation. We show in Section 3 that the method defined by (7)-(8) is an order optimal regularization method for suitable parameters α. Moreover, we show that both fractional methods defined by (7)-(8) and (10) are order optimal when used with the discrepancy principle. A discussion on advantages and disadvantages of these fractional methods concludes the section. Section 4 contains a few illustrative numerical examples, and concluding remarks can be found in Section 5.

2 Regularization methods and filter factors

This section reviews definitions and properties of regularization methods; see, e.g., [2, 7] for further details. A *regularization method* for $A^\dagger$ is a family of operators

$$\{R_\mu\}_{\mu>0}, \quad R_\mu : Y \to X$$

with the following properties: There is a mapping $\mu : \mathbb{R}_+ \times Y \to \mathbb{R}_+$ such that for all $b \in \mathcal{D}(A^\dagger)$ and all $b^\delta \in Y$ with $\|b - b^\delta\|_Y \leq \delta$, it holds

$$\lim_{\delta \searrow 0} R_{\mu(\delta, b^\delta)} b^\delta = A^\dagger b.$$

Here μ is a regularization parameter.

The quality of a regularization method is determined by the asymptotics of $\|A^\dagger b - R_\mu b^\delta\|_X$ as $\delta \searrow 0$. Convergence rates can only be achieved under additional assumptions on the solution. For our analysis, we assume a Hölder-type smoothness assumption, i.e., that the minimal norm solution $x^\dagger$ of the error-free problem (1) satisfies a smoothness condition of the form

$$x^\dagger \in \operatorname{range}((A^*A)^{\nu/2}) \quad \text{with} \quad \|x^\dagger\|_\nu := \Big(\sum_{n\geq 1} \sigma_n^{-2\nu} |\langle x^\dagger, u_n\rangle|^2\Big)^{1/2} \leq \rho \tag{11}$$

for some constant ρ. Here $(\sigma_n; u_n, v_n)_{n\geq 1}$ is the singular system of the operator A. A regularization method is said to be *order optimal* if there is a constant c independent of δ and ρ such that

$$\|x^\dagger - R_\mu b^\delta\|_X \leq c\,\delta^{\frac{\nu}{\nu+1}} \cdot \rho^{\frac{1}{\nu+1}}.$$

It is well known that Tikhonov regularization in standard form is an order optimal method, see, e.g., [2].

Generalized Tikhonov regularization is obtained by replacing the penalty term in (4) by $\|Bx\|_X^2$, where $B : \mathcal{N}(A)^\perp \to X$ is an operator whose domain $\mathcal{D}(B)$ is dense in $\mathcal{N}(A)^\perp$ and $(B^*B)^{-1} : \mathcal{N}(A)^\perp \to X$ is continuous. Here $\mathcal{N}(A)^\perp$ denotes the orthogonal complement of the null space of A. The associated functional is

$$J_{\mu,B}(x) := \|Ax - b^\delta\|_Y^2 + \mu\|Bx\|_X^2. \tag{12}$$

Certain conditions on the operator B allow for results on optimality and order optimality of generalized Tikhonov regularization; see Louis [7]. Proposition 3.1 below shows the equivalence of generalized Tikhonov regularization with a special operator B and fractional Tikhonov regularization (8).

Filter factors provide insight into the properties of regularization methods. Let the linear compact operator A have the singular system $(\sigma_n; u_n, v_n)_{n\geq 1}$. We replace the Moore–Penrose generalized inverse of A by an operator R_μ defined by

$$R_\mu b^\delta := \sum_{\sigma_n > 0} F_\mu(\sigma_n)\sigma_n^{-1}\langle b^\delta, v_n\rangle u_n. \tag{13}$$

The real-valued function F_μ is referred to as a filter function and its values $F_\mu(\sigma_n)$ as filter factors; $\mu > 0$ is a regularization parameter. Thus, $R_\mu b^\delta$ furnishes an approximation of $x^\dagger$. For example, Tikhonov regularization in standard form can be characterized by the filter function

$$F_\mu^{\mathrm{Tikh}}(\sigma) = \frac{\sigma^2}{\sigma^2 + \mu}. \tag{14}$$

That is, the minimizer of (4) can also be computed as

$$x_\mu^\delta = \sum_{\sigma_n > 0} \frac{\sigma_n}{\sigma_n^2 + \mu}\langle b^\delta, v_n\rangle u_n.$$

as well as by (5). The filter function associated with the fractional Tikhonov regularization method (7)-(8) is given by

$$\widetilde{F}_{\mu,\alpha}(\sigma) = \frac{\sigma^{\alpha+1}}{\sigma^{\alpha+1} + \mu} \tag{15}$$

and gives the associated approximation

$$\widetilde{x}_{\mu}^{\delta} = \sum_{\sigma_n > 0} \frac{\sigma_n^{\alpha}}{\sigma_n^{\alpha+1} + \mu} \langle b^{\delta}, v_n \rangle u_n \tag{16}$$

of $x^{\dagger}$. This expression is provided in [4] with slightly different notation.

The fractional Tikhonov method (10) can be written in terms of a filter function in a similar fashion. We have

$$\widehat{F}_{\mu,\alpha}(\sigma) := (F_{\mu}^{\text{Tikh}}(\sigma))^{\alpha} = \left(\frac{\sigma^2}{\sigma^2 + \mu} \right)^{\alpha}. \tag{17}$$

The corresponding approximation of $x^{\dagger}$ is given by

$$\widehat{x}_{\mu}^{\delta} = \sum_{\sigma_n > 0} \left(\frac{\sigma_n^{2-1/\alpha}}{\sigma_n^2 + \mu} \right)^{\alpha} \langle b^{\delta}, v_n \rangle u_n. \tag{18}$$

3 Order optimality of fractional Tikhonov methods

We first discuss the order optimality of the fractional regularization methods (7)-(8) and (10).

Proposition 3.1. *Let $A : X \to Y$ be a linear compact operator between Hilbert spaces X and Y. Let $x^{\dagger} := A^{\dagger} b$ satisfy $\|x^{\dagger}\|_{\nu} \le \rho$, where the norm is defined by* (11)*. Then for all exponents $\alpha \ge 0$ and the parameter choice rule*

$$\mu = C \left(\frac{\delta}{\rho} \right)^{(\alpha+1)/(2(\nu+1))}, \tag{19}$$

the fractional Tikhonov method (7)-(8) *is order optimal. Here $C > 0$ is a constant independent of δ and ρ.*

Proof. Solutions of the minimization problems associated with the functionals (8) and (12) can be determined from the associated normal equations. For generalized Tikhonov regularization (12), we obtain with $B^* B = (A^* A)^{-\eta}$ the normal equations

$$(A^* A + \mu (A^* A)^{-\eta}) x_{\mu}^{\delta} = A^* b^{\delta}, \tag{20}$$

where $(A^* A)^{-1}$ is replaced by the Moore–Penrose pseudoinverse if A is not of full rank. Louis [7, Satz 4.2.3] establishes that this method is order optimal for $\eta \ge 1/2$. A comparison with (9) shows that the fractional Tikhonov method (8) is order optimal for $\alpha \ge 0$. □

While the method (8) is order optimal for all $\alpha \geq 0$, this is not the case for the fractional Tikhonov method (10). We have the following result.

Proposition 3.2. *[6, Proposition 3.2] Let $A : X \to Y$ be a compact operator with singular system $(\sigma_n, u_n, v_n)_{n\geq 0}$, and let $x^\dagger := A^\dagger b$ satisfy $\|x^\dagger\|_\nu \leq \rho$ for some constant ρ and the ν-norm defined by (11). Then for $\alpha \in (1/2, 1]$, the fractional Tikhonov method defined (10) is order optimal with the parameter choice rule*

$$\mu = C\left(\frac{\delta}{\rho}\right)^{1/2(\nu+1)}$$

for all $0 < \nu < 2$. Here C is a positive constant independent of δ and ρ.

Klann and Ramlau [6, Theorem 4.4] show that after appropriate presmoothing of the error-contaminated data b^δ, fractional powers $0 < \alpha \leq 1/2$ together with a suitable choice of the regularization parameter μ yield quasi-optimal convergence rates.

The above approaches to determine μ generally are not very useful for the solution of specific problems. When an accurate estimate of the norm of the error in the data $\|b^\delta - b\|_Y$ is known, the *discrepancy principle*, discussed, e.g., in [2, 8], can be applied to determine a suitable value of μ. The idea is to choose the value of μ so that the residual is approximately of the same norm as the error in the data b^δ. There are several slightly different formulations of the discrepancy principle. Here we will choose $\mu = \mu(\delta, b^\delta)$ such that

$$\|Ax_\mu - b^\delta\| = \tau\delta, \tag{21}$$

where $\tau > 1$ is a user-supplied constant independent of δ. This is a nonlinear equation for μ. Its solution can be calculated by finding the positive zero of

$$G_\alpha(\mu) := \sum_{n\in\mathbb{N}} (1 - F_{\mu,\alpha}(\sigma))^2 \langle b^\delta, v_n\rangle^2 - (\tau\delta)^2, \tag{22}$$

for example with Newton's method; see, e.g., [4] for further details.

Convergence of regularized approximate solutions determined by filtered regularization methods using the discrepancy principle has been analyzed by Louis [7], who shows the following result (with slightly different notation):

Theorem 3.1. *[7, Theorem 3.5.2] Let $b \in \operatorname{range}(A)$ and $\|b - b^\delta\|_Y \leq \delta$. For all $\sigma \in (0, \sigma_1]$, let $\mu \mapsto |1 - F_{\mu,\alpha}|$ be continuous and monotonically increasing, and assume that for $0 \leq \nu \leq 2$,*

$$\sup_{0<\sigma\leq\sigma_1} \sigma^{-1}|F_{\mu,\alpha}(\sigma)| \leq c\mu^{-\alpha/2} \quad \text{and} \quad \sup_{0<\sigma\leq\sigma_1} |1 - F_{\mu,\alpha}(\sigma)|\sigma^\nu \leq c_\nu \mu^{\alpha\nu/2}$$

for a constant c independent of δ and ν, and a constant c_ν independent of δ. Let $\mu = \mu(\delta, b^\delta)$ be determined by (21). Then

$$R_\mu b^\delta \to A^\dagger b \;\; \text{for} \;\; \delta \searrow 0.$$

Assume that $x^\dagger$ can be written as $x^\dagger = (A^*A)^{\nu/2}h$ for some h such that $\|h\| \le \rho$ and some $0 < \nu \le 1$. Then there is a constant d_ν independent of b, δ, ρ such that

$$\|A^\dagger b - R_\mu b^\delta\| \le d_\nu \delta^{\nu/(\nu+1)} \cdot \rho^{1/(\nu+1)}.$$

Both fractional methods (7)-(8) and (10) satisfy the conditions of the above theorem. This allows us to show that these methods are order optimal with respect to the discrepancy principle.

Corollary 3.1. *Let $A : X \to Y$ be a linear compact operator between Hilbert spaces X and Y. Let $x^\dagger := A^\dagger b$ satisfy $\|x^\dagger\|_\nu \le \rho$. Then for all exponents $\alpha \ge 0$, the fractional Tikhonov method* (7)-(8) *is order optimal with the regularization parameter μ determined by the discrepancy principle* (21)*.*

Proof. Using the equivalence between (9) and (20), the proof follows from results in [7], specifically from Chapter 3.5 together with Satz 4.2.2 and Satz 4.2.3. □

Remark 3.1. *It might appear appealing to substitute the standard norm in* (21) *by the weighted norm from* (8)*. Then with $W = (A^*A)^{(\alpha-1)/2}$,*

$$\|A\widetilde{x}_\mu - b^\delta\|_W^2 = \sum_n (1 - \widetilde{F}_\mu(\sigma_n))^2 \sigma_n^{\alpha-1} \langle b^\delta, v_n\rangle^2.$$

However, since $\lim_{\sigma_n \to 0} \widetilde{F}_\mu(\sigma_n) = 0$, the sum will not converge since for large n the inner products $\langle b^\delta, v_n\rangle$ are dominated by the error in b^δ and do not converge to zero. In a discrete setting, the residual will be very large due to noise amplification, and equation (21) *is not guaranteed to have a solution. Hence, the weighted residual norm is in general not useful.*

Corollary 3.2. *Let $A : X \to Y$ be a linear compact operator between Hilbert spaces X and Y. Let $x^\dagger := A^\dagger b$ satisfy $\|x^\dagger\|_\nu \le \rho$. Then for all exponents $\alpha \in (1/2, 1]$, the fractional Tikhonov method of* (10) *is order optimal with the regularization parameter μ given by the discrepancy principle* (21)*.*

Proof. It is shown in [6, Lemma 3.1 and Proposition 3.2] that the conditions of Theorem 3.1 hold. □

Approximations of $x^\dagger$ determined by fractional Tikhonov regularization typically are closer to $x^\dagger$ in the X-norm than approximations obtained with Tikhonov regularization in standard form; see [4] for computed examples. However, a smaller error does not always correspond to a more pleasing approximation of $x^\dagger$, because the fractional Tikhonov approximation may be more oscillatory than the approximation determined by Tikhonov regularization in standard form. We would like to elucidate in which situations fractional methods yield more pleasing approximations. The following lemma is helpful. A similar result has been shown in [4].

Lemma 3.1. *The mappings $\mu \mapsto F_{\mu,\alpha}(\sigma)$ and $\alpha \mapsto F_{\mu,\alpha}(\sigma)$ are continuous and monotonically decreasing for $\mu > 0$ and α in an interval $\underline{\alpha} < \alpha < \overline{\alpha}$. Let $\mu = \mu(\alpha)$ be determined by the discrepancy principle* (21)*. Then $\frac{d\mu(\alpha)}{d\alpha} < 0$.*

Proof. We can write (22) in the form $G(\alpha, \mu(\alpha)) = 0$. Since G is differentiable, we have

$$\frac{dG}{d\mu} = \sum_{\sigma_n > 0} 2(1 - F_{\mu,\alpha}(\sigma_n)) \cdot (-1) \cdot \frac{dF_{\mu,\alpha}}{d\mu} \cdot \langle b^\delta, v_n \rangle^2 \quad > 0,$$

because $1 - F_{\mu,\alpha}(\sigma) > 0$ and $\frac{dF_{\mu,\alpha}}{d\mu} < 0$. Analogously, one finds that $\frac{dG}{d\alpha} > 0$. Hence, by the implicit function theorem,

$$\frac{d\mu}{d\alpha} = -\left(\frac{dG}{d\mu}\right)^{-1} \frac{dG}{d\alpha} < 0.$$

$\square$

An immediate consequence of the above lemma is that decreasing α results in an increase of the regularization parameter μ. It is therefore inappropriate to compare fractional methods with the standard Tikhonov filter using the same regularization parameter.

We are now in position to have a closer look at the computed approximations. Again we will make use of the explicit representation of the solution in terms of the singular system of A. Let

$$\epsilon = b^\delta - b.$$

Since

$$\sigma_n \langle x^\dagger, u_n \rangle = \langle x^\dagger, A^* v_n \rangle = \langle b, v_n \rangle,$$

cf. [2], and

$$\langle b^\delta, v_n \rangle = \langle b, v_n \rangle + \langle \epsilon, v_n \rangle,$$

the approximation error $e(\delta, \alpha, \mu) := x^\dagger - R_\mu b^\delta$ is given by

$$e(\delta, \alpha, \mu) = \sum_{\sigma_n > 0} (1 - F_{\mu,\alpha}(\sigma_n)) \langle x^\dagger, u_n \rangle u_n + \sum_{\sigma_n > 0} F_{\mu,\alpha}(\sigma_n) \frac{1}{\sigma_n} \langle -\epsilon, v_n \rangle u_n. \quad (23)$$

Let ϵ be fixed. The performance of the reconstruction is then determined by the positive coefficients $1 - F_{\mu,\alpha}(\sigma)$ and $F_{\mu,\alpha}(\sigma)$. One immediately sees that the filter has to achieve two contradicting properties: $F_{\mu,\alpha}(\sigma)$ should be close to one to give a small deviation of the reconstruction from $x^\dagger$, and also $F_{\mu,\alpha}(\sigma)$ should be close to zero in order to effectively reducing propagation of the error ϵ into the computed approximation.

It is not obvious from (23) in which situations letting $\alpha < 1$ improves the quality of the computed approximation of $x^\dagger$. We can shed some light on this by studying the derivative $\frac{d}{d\alpha} F_{\mu,\alpha}(\sigma)$. We first consider the filter function (15). Since μ depends on α, we get

$$\frac{d}{d\alpha} \widetilde{F}_{\mu,\alpha}(\sigma) = -\frac{d}{d\alpha}\left(1 - \widetilde{F}_{\mu,\alpha}(\sigma)\right) = h(\sigma, \alpha, \mu(\alpha)) \left(\ln \sigma - \frac{\mu'(\alpha)}{\mu(\alpha)}\right), \quad (24)$$

where $h(\sigma, \alpha, \mu(\alpha))$ is a positive function. The sign of the derivative is determined by the factor $\ln \sigma - \frac{\mu'(\alpha)}{\mu(\alpha)}$. When α and the error norm δ are fixed, so is μ, and the sign only depends on σ. By Lemma 3.1, $\mu'(\alpha) < 0$. Therefore, the derivative (24) changes sign at some $0 < \widetilde{\sigma}_0 < 1$. Only for n with $\sigma_n < \widetilde{\sigma}_0$, the coefficient of $\langle x^\dagger, u_n\rangle$ in (23) will be reduced by decreasing α, since then $\frac{d}{d\alpha}(1 - \widetilde{F}_{\mu,\alpha}(\sigma_n)) > 0$. Hence, the coefficient of $\langle x^\dagger, u_n\rangle$ increases. The opposite holds true for the coefficients of the terms associated with the propagated error. Whereas for large σ_n the propagated error is damped, it is amplified for all $\sigma_n < \widetilde{\sigma}_0$.

The result for the fractional filter (17) is analogous. Similarly to (24), one has

$$\begin{aligned}\frac{d}{d\alpha}\widehat{F}_{\mu,\alpha}(\sigma) &= -\frac{d}{d\alpha}\left(1 - \widehat{F}_{\mu,\alpha}(\sigma)\right)\\ &= \widehat{h}(\sigma,\alpha,\mu(\alpha))\left(-\ln\left(\frac{\sigma^2+\mu(\alpha)}{\sigma^2}\right) - \alpha\frac{\mu'(\alpha)}{\sigma^2+\mu(\alpha)}\right) \qquad (25)\end{aligned}$$

with $\widehat{h}(\sigma,\alpha,\mu(\alpha)) > 0$. The logarithm is positive and $\mu'(\alpha) < 0$. Therefore, the sign of (25) changes at some $\sigma = \widehat{\sigma}_0 > 0$. Hence, the above discussion also applies to this filter function. However, it is not clear whether the operator A has singular values that satisfy $\sigma_n > \widehat{\sigma}_0$. If this is not the case, then decreasing α will result in error amplification in all components of the computed approximate solution.

Although it is an open problem how to determine a value of α that yields the best approximation of $x^\dagger$, we can identify two situation in which fractional Tikhonov methods outperform standard Tikhonov regularization (3)-(4):

a) the problem is severely ill-posed, i.e., the singular values of A decrease rapidly to zero, and

b) the error in b^δ is concentrated to low frequencies.

In case the problem is severely ill-posed, $\widetilde{\sigma}_0$ and $\widehat{\sigma}_0$ are likely to be large enough for the propagated error to be damped. A slight loss in accuracy of terms in (16) and (18) associated with large singular values is typically acceptable, since they are much larger than the error and therefore usually are recovered quite accurately. On the other hand, if there is only little error in the high frequency components in (16) and (18), the amplification of the error in b^δ is largely avoided, while the reconstruction is improved. In other cases, both fractional methods do not perform significantly better than Tikhonov regularization in standard form. The reason for this can again be found in the dependency of the filter factors $F_{\mu,\alpha}(\sigma)$ on the parameters α and μ. By decreasing α, the $F_{\mu,\alpha}(\sigma)$ increase. At the same time, decreasing α leads to increasing regularization parameter μ as shown in Lemma 3.1. From the definition of the filter factors (15) and (17), respectively, one sees that this leads to decreasing values of the filter factors. Hence, both effects cancel each others out to some extend. Although α is decreased below one, the filter factors corresponding to larger singular values stay almost constant. The following section provide some illustrative computed examples.

4 Numerical examples

We illustrate the theory developed in the previous section with two examples, the first of which is a severely ill-posed Fredholm integral equation of the first kind given by

$$b(s) = [A_1 x](s) = \int_0^1 \sqrt{s^2+t^2} f(t)dt, \quad 0 \le s \le 1, \tag{26}$$

with error-free data $b(s) = \frac{1}{3}\left((1+s^2)^{3/2} - s^3\right)$ and solution $x^\dagger(t) = t$. This equation was first introduced by Fox and Goodwin, cf. [1]. Numerically, the singular values decrease exponentially until they stagnate around attainable computational precision.

The second example is the mildly ill-posed Volterra integral equation of the first kind

$$b(s) = [A_2 x](s) = \int_0^s f(t)dt, \quad 0 \le s \le 1, \tag{27}$$

with error-free data

$$b_1(s) = \begin{cases} -s & 0 \le s \le 0.5, \\ s-1 & 0.5 < s \le 1, \end{cases}$$

and solution

$$x_1^\dagger(t) = \begin{cases} -1 & 0 \le t \le 0.5, \\ 1 & 0.5 < t \le 1. \end{cases}$$

The same example was used in [6]. The coefficients $\langle x_1^\dagger, u_n\rangle$ decrease slowly to zero. In order to demonstrate that the performance of the fractional Tikhonov methods mainly depends on properties of the operator, we also used a data function that gives an alternative solution $x_2^\dagger$, which is designed so that $\langle x_2^\dagger, u_n\rangle = \mathcal{O}(\exp(-n))$. After discretizing the operator A_2 in (27) we computed the SVD of the resulting matrix and hence the u_n were available. The singular system $\{\sigma_n; u_n, v_n\}_{n\ge 1}$ of A_2 (without discretization) is given in [7]. We found the quality of the computed solution to be the same for both problems (26) and (27). Numerically, the implementation of (26) in the Regularization Toolbox [3] has been used. The integration problem (27) has been discretized simply with the trapezoidal rule. In all experiments shown, we used 100 discretization points and equipped both the domain and range of the discretized operators with the Euclidean vector norm.

All plots compare approximate solutions obtained by the two fractional methods (7)-(8) and (10) with the approximate solution determined by Tikhonov regularization in standard form (4). Also the desired solution of the error-free problem $x^\dagger$ is shown. For all approximate solutions, the regularization parameter is determined by the discrepancy principle (21). Figure 1 shows that, for the severely ill-posed problem (26) and an error in b^δ made up of 5% relative Gaussian noise, the fractional method (7)-(8) determines a much more accurate approximation of $x^\dagger$ than Tikhonov regularization in standard form, whereas the approximate solution determined by the fractional method (10) is only slightly more accurate than the approximate solution determined by standard Tikhonov regularization. For the fractional methods, we calculated solutions for several $0 \leq \alpha \leq 1$ and plotted the ones that gave the best approximation of $x^\dagger$. The same technique, when applied to the problem (27), again with 5% relative Gaussian noise in the data b^δ, lead to α-values one or very close to one. Since then the fractional methods are very close to Tikhonov regularization in standard form, we plotted solutions for a fixed, smaller α instead. Figure 2 shows that, for both exact solutions $x_1^\dagger$ and $x_2^\dagger$, the approximate solutions determined by the fractional methods are not more accurate than the one determined by standard Tikhonov regularization. Indeed, the amplification of the error in the data is clearly visible. For problems which have a degree of ill-posedness between the ones shown here, we observed that the quality of the computed approximate solution strongly depends on the realization of the noise. That is, inverting several noisy data sets where all parameters, including the noise level, were kept constant, the fractional methods sometimes gave approximate solutions of considerably higher quality than Tikhonov regularization, while in other cases there was no improvement in quality.

So far the data was perturbed by white Gaussian noise. Figure 4 shows results obtained with low-frequency noise. An example of this kind of noise in comparison with white noise is shown in Figure 3. The fractional methods clearly give more accurate approximations of $x^\dagger$ than Tikhonov regularization in standard form for low-frequency noise.

To further show the different behavior of the methods in our comparison in the settings introduced above, we include tables in which we give regularization parameters and approximation errors relative to those obtained with Tikhonov regularization in standard form,

$$\widetilde{\mathrm{re}} = \|\tilde{x}^\delta_{\mu(\alpha)} - x^*\| / \|x^\delta_{\mu(1)} - x^*\| \tag{28}$$

for the fractional method (7)-(8) and

$$\widehat{\mathrm{re}} = \|\tilde{x}^\delta_{\mu(\alpha)} - x^*\| / \|x^\delta_{\mu(1)} - x^*\| \tag{29}$$

for the method (10) for several values of α. All errors are averages over 20 experiments with different error-realizations. Table 1 shows results for the Fox–Goodwin problem (26). In agreement with Figure 1, the fractional method (7)-(8) performs the best. For the problem (27) with Gaussian white noise, the error in the approximate solutions determined by the fractional methods is only slightly smaller than the one obtained with Tikhonov regularization in standard form, as shown in Table 2. However, using the same problem with low-frequency error instead of white Gaussian error, the fractional methods yield a much better approximations of $x^\dagger$ than Tikhonov regularization in standard form; see Table 3.

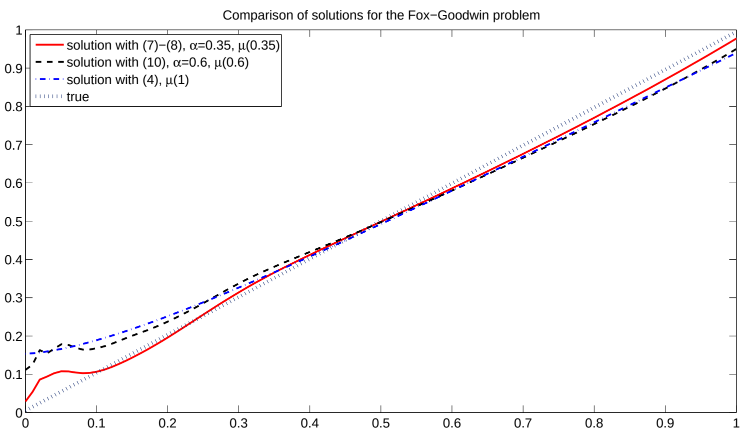

Figure 1: Comparison of solutions for the severely ill-posed Fox-Goodwin problem (26) with 5% relative white Gaussian noise in the data, μ according to (21), $\tau = 1.1$. For the fractional methods the solutions with lowest reconstruction error are shown. The solution for the method (7)-(8) is plotted with the solid line. For this type of problems it is to be preferred over the other two methods. Those are the method (10) (dashed) an Tikhonov regularization in standard form (dash-dotted).

α	0.05	0.1	0.3	0.5	0.6	0.7	0.9	1
$\widetilde{\mu}$	6.1e-3	5.9e-3	5.2e-3	4.3e-3	3.8e-3	3.3e-3	2.4e-3	2.0e-3
$\widehat{\mu}$	1.1e-1	3.9e-2	8.2e-3	4.4e-3	3.5e-3	2.9e-3	2.2e-3	2.0e-3
$\widetilde{\mathrm{re}}$	10.3	5.2	1.5	0.60	0.61	0.72	0.92	1
$\widehat{\mathrm{re}}$	2.4e15	4.7e13	1e7	4.2	0.93	0.91	0.97	1

Table 1: Regularization parameter and relative reconstruction error for both fractional filters, tilde standing for (7)-(8), hat for (17); and the Fox-Goodwin problem (26). In both cases μ grows monotonically with decreasing α. The reconstruction errors (28) and (29), respectively, are shown in the two bottom rows. For the method (7)-(8), there is a minimum clearly below one. Hence, the reconstructions are significantly improved. Since for $\alpha < 0.5$ the filter (17) is not regularizing anymore, the reconstruction error explodes.

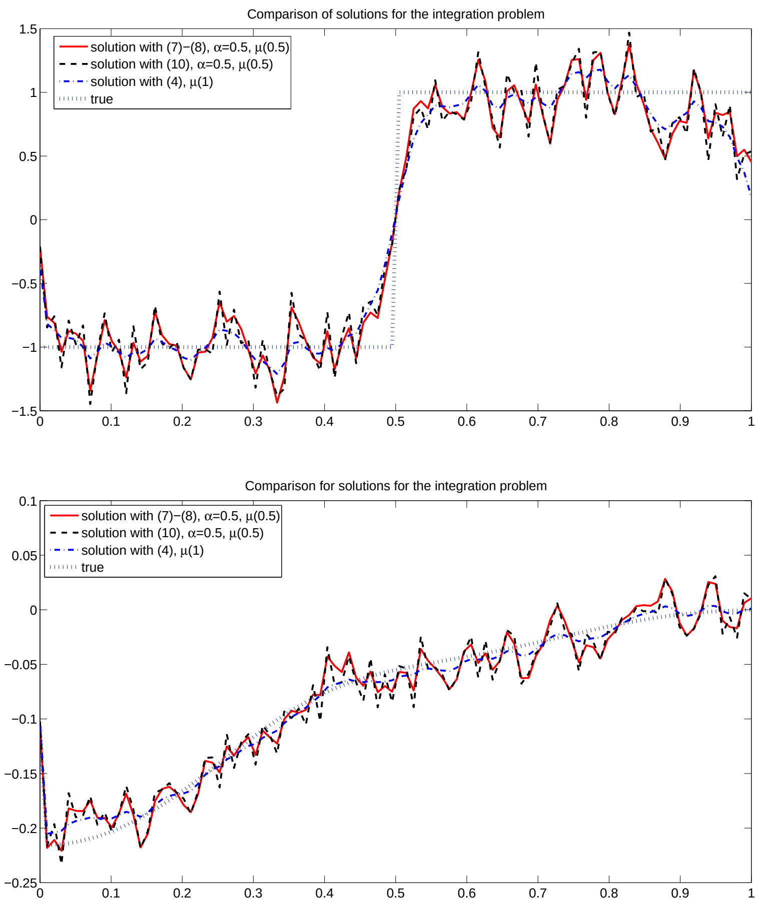

Figure 2: Comparison of solutions for the mildly ill-posed integration problem (27) with 5% relative white Gaussian noise in the data, μ according to (21), $\tau = 1.1$. Upper plot: discontinuous solution, lower plot: smooth solution. In this case the fractional methods (7)-(8) and (10) do not perform better than Tikhonov regularization in standard form. On the contrary, the noise is amplified even more. The lowest reconstruction error was achieved for $\alpha = 0.95$ and $\alpha = 1$, respectively.

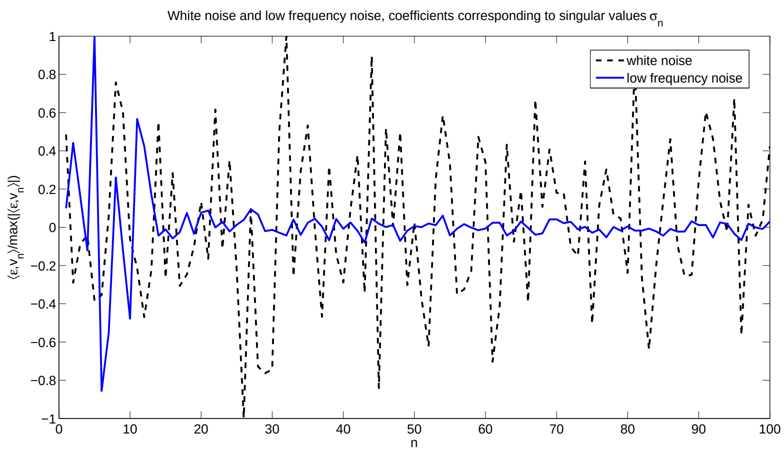

Figure 3: Comparison of typical random draws of white noise and low frequency noise w.r.t. the singular values. For white noise, the coefficients are equally distributed over all singular values. The low frequency noise decreases with growing n.

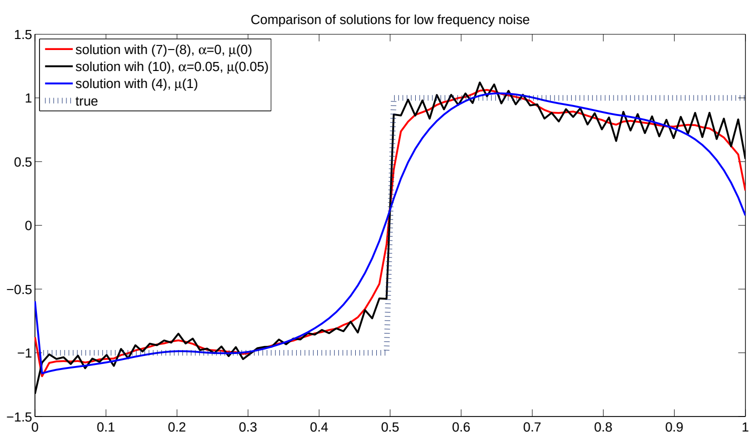

Figure 4: Comparison of solutions for the mildly ill-posed integration problem, 5% relative low frequency noise (cf. Figure 3) in the data, μ according to (21), $\tau = 1.1$. The solution of the fractional methods (7)-(8) (solid) and (10) (dashed) with appropriate α approximate the discontinuity much better than the results of Tikhonov regularization in standard form (dash-dotted).

α	0.05	0.1	0.3	0.5	0.6	0.7	0.9	1
$\widetilde{\mu}$	4.7e-3	4.3e-3	2.7e-3	1.6e-3	1.3e-3	1.0e-3	0.6e-3	0.4e-3
$\widehat{\mu}$	3.4e-2	1.1e-2	2.3e-3	1.1e-3	0.8e-3	0.7e-3	0.5e-3	0.4e-3
$\widetilde{\mathrm{re}}$	2.7	2.3	1.4	1.07	1.02	0.991	0.990	1
$\widehat{\mathrm{re}}$	21.8	12.5	2.3	1.13	1.03	1.007	0.998	1

Table 2: Regularization parameter and relative reconstruction error for both fractional filters and the integration problem (27). In both cases μ grows monotonically with decreasing α. The reconstruction errors (28) and (29), respectively, grow nearly monotonically, only for α close to one it is slightly below one, i.e., the fractional methods give a slightly lower residual than Tikhonov regularization in standard form.

α	0.05	0.1	0.3	0.5	0.6	0.7	0.9	1
$\widetilde{\mu}$	6.0e-3	5.5e-3	3.8e-3	2.6e-3	2.1e-3	1.7e-3	1.1e-3	0.9e-3
$\widehat{\mu}$	3.9e-2	1.4e-2	3.5e-3	1.9e-3	1.5e-3	1.3e-3	1.0e-3	0.9e-3
$\widetilde{\mathrm{re}}$	0.62	0.65	0.75	0.84	0.88	0.91	0.97	1
$\widehat{\mathrm{re}}$	0.54	0.60	0.81	0.91	0.94	0.96	0.99	1

Table 3: Regularization parameter and relative reconstruction error for both filters and the integration problem (27) in presence of low frequency noise (cf. Figure 3). The reconstruction errors (28) and (29), respectively, are shown in the two bottom rows. Both fractional filters give a much better result than Tikhonov regularization in standard form.

5 Conclusion

We have further investigated the fractional Tikhonov method of [4, 7], (7)-(8) and the fractional Tikhonov method [6], (10), and showed that the method of Louis is of optimal order for a certain interval of parameters α with an appropriate choice of the regularization parameter. Moreover, we demonstrated that both methods are of optimal order with the discrepancy principle. Two situations in which the fractional methods are significantly better than Tikhonov regularization in standard form are illustrated with numerical examples.

References

[1] C. T. H. Baker. *The Numerical Treatment of Integral Equations.* Clarendon Press, Oxford, 1977.

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

[3] P. C. Hansen. Regularization tools version 4.0 for Matlab 7.3. Numer. Algorithms, 46:189-194, 2007.

[4] M. E. Hochstenbach and L. Reichel. Fractional Tikhonov regularization for linear discrete ill-posed problems. *BIT Numerical Mathematics*, 51:197–215, 2011.

[5] M. E. Hochstenbach, S. Noschese, and L. Reichel. Fractional regularization matrices for linear discrete ill-posed problems. *J. Eng. Math.*, in press.

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

[7] A. K. Louis. *Inverse und schlecht gestellte Probleme.* Stuttgart: Teubner, 1989.

[8] V. A. Morozov. *Methods for Solving Incorrectly Posed Problems.* New York: Springer, 1984.

[9] A. N. Tikhonov and V. Y. Arsenin. *Solutions of Ill-Posed Problems.* New York: Wiley, 1977.

Technical Reports of the Doctoral Program

"Computational Mathematics"

2014

2014-01 E. Pilgerstorfer, B. Jüttler: *Bounding the Influence of Domain Parameterization and Knot Spacing on Numerical Stability in Isogeometric Analysis* February 2014. Eds.: B. Jüttler, P. Paule

2014-02 T. Takacs, B. Jüttler, O. Scherzer: *Derivatives of Isogeometric Functions on Rational Patches* February 2014. Eds.: B. Jüttler, P. Paule

2014-03 M.T. Khan: *On the Soundness of the Translation of MiniMaple to Why3ML* February 2014. Eds.: W. Schreiner, F. Winkler

2014-04 G. Kiss, C. Giannelli, U. Zore, B. Jüttler, D. Großmann, J. Barne: *Adaptive CAD model (re–)construction with THB–splines* March 2014. Eds.: M. Kauers, J. Schicho

2014-05 R. Bleyer, R. Ramlau: *An Efficient Algorithm for Solving the dbl-RTLS Problem* March 2014. Eds.: E. Klann, V. Pillwein

2014-06 D. Gerth, E. Klann, R. Ramlau, L. Reichel: *On Fractional Tikhonov Regularization* April 2014. Eds.: W. Zulehner, U. Langer

2013

2013-01 U. Langer, M. Wolfmayr: *Multiharmonic Finite Element Analysis of a Time-Periodic Parabolic Optimal Control Problem* January 2013. Eds.: W. Zulehner, R. Ramlau

2013-02 M.T. Khan: *Translation of* MiniMaple *to Why3ML* February 2013. Eds.: W. Schreiner, F. Winkler

2013-03 J. Kraus, M. Wolfmayr: *On the robustness and optimality of algebraic multilevel methods for reaction-diffusion type problems* March 2013. Eds.: U. Langer, V. Pillwein

2013-04 H. Rahkooy, Z. Zafeirakopoulos: *On Computing Elimination Ideals Using Resultants with Applications to Gröbner Bases* May 2013. Eds.: B. Buchberger, M. Kauers

2013-05 G. Grasegger: *A procedure for solving autonomous AODEs* June 2013. Eds.: F. Winkler, M. Kauers

2013-06 M.T. Khan *On the Formal Verification of Maple Programs* June 2013. Eds.: W. Schreiner, F. Winkler

2013-07 P. Gangl, U. Langer: *Topology Optimization of Electric Machines based on Topological Sensitivity Analysis* August 2013. Eds.: R. Ramlau, V. Pillwein

2013-08 D. Gerth, R. Ramlau: *A stochastic convergence analysis for Tikhonov regularization with sparsity constraints* October 2013. Eds.: U. Langer, W. Zulehner

2013-09 W. Krendl, V. Simoncini, W. Zulehner: *Efficient preconditioning for an optimal control problem with the time-periodic Stokes equations* November 2013. Eds.: U. Langer, V. Pillwein

The complete list since 2009 can be found at
`https://www.dk-compmath.jku.at/publications/`

Doctoral Program
"Computational Mathematics"

Director:

Prof. Dr. Peter Paule
Research Institute for Symbolic Computation

Deputy Director:

Prof. Dr. Bert Jüttler
Institute of Applied Geometry

Address:

Johannes Kepler University Linz
Doctoral Program "Computational Mathematics"
Altenbergerstr. 69
A-4040 Linz
Austria
Tel.: ++43 732-2468-6840

E-Mail:

office@dk-compmath.jku.at

Homepage:

http://www.dk-compmath.jku.at