

# ON FRACTIONAL ASYMPTOTICAL REGULARIZATION OF LINEAR ILL-POSED PROBLEMS IN HILBERT SPACES

Ye Zhang <sup>1</sup>, Bernd Hofmann <sup>2</sup>

## Abstract

In this paper, we study a fractional-order variant of the asymptotical regularization method, called *Fractional Asymptotical Regularization (FAR)*, for solving linear ill-posed operator equations in a Hilbert space setting. We assign the method to the general linear regularization schema and prove that under certain smoothness assumptions, FAR with fractional order in the range  $(1, 2)$  yields an acceleration with respect to comparable order optimal regularization methods. Based on the one-step Adams-Moulton method, a novel iterative regularization scheme is developed for the numerical realization of FAR. Two numerical examples are given to show the accuracy and the acceleration effect of FAR.

*MSC 2010:* Primary 47A52; Secondary 26A33, 65J20, 65F22, 65R30,

*Key Words and Phrases:* Linear ill-posed operator equation, asymptotical regularization, fractional derivatives, stopping rules, source conditions, convergence rates, acceleration

## 1. Introduction

It belongs to the main areas of competence of the journal “Fractional Calculus & Applied Analysis” to publish papers that handle initial-boundary value problems and abstract initial-value problems in Hilbert spaces for time-fractional diffusion and wave equations, partially also with respect to inverse and ill-posed problems. In this context, we refer for example to the papers [14, 17, 18, 19] by M. YAMAMOTO, Y. LUCHKO, and coauthors. The behavior of solutions to such equations plays an important role for finding stable approximations to inverse problems aimed at characterizing parameter functions of the underlying physical processes. As the papers

---

mentioned above show, error estimates and explicit representations of solutions to the fractional equations can also be used to generate new continuous regularization methods. In the present paper, we try to complement this research direction by extending the well-known method of asymptotical regularization (Showalter's method) in Hilbert spaces to a fractional version using the left-side Caputo fractional derivative. In particular, we are going to derive regularization properties by a stringent analysis and to show at least by means of case studies an acceleration effect based on the fractional derivatives.

Let  $A$  be a compact linear operator acting between two infinite dimensional Hilbert spaces  $\mathcal{X}$  and  $\mathcal{Y}$  such that the range  $\mathcal{R}(A)$  of  $A$  is an infinite dimensional subspace of  $\mathcal{Y}$ . It is well-known that then  $\mathcal{R}(A)$  is non-closed, i.e.  $\overline{\mathcal{R}(A)}^{\mathcal{Y}} \neq \mathcal{R}(A)$ , and the linear operator equation

$$Ax = y, \quad (1.1)$$

is *ill-posed* and requires some kind of *regularization*. This is typical for operator equations (1.1) which are models for *linear inverse problems* (cf., e.g., [9] or [10, § 2] and references therein). For simplicity, we denote by  $\langle \cdot, \cdot \rangle$  and  $\|\cdot\|$  in the sequel the inner products and norms for both Hilbert spaces  $\mathcal{X}$  and  $\mathcal{Y}$ . In this paper, we are interested in stable approaches for solving (1.1) that can be realized by fast iterative algorithms. The corresponding stable approximate solutions should be based on noisy data  $y^\delta$  of the exact right-hand side  $y = Ax^\dagger$ , where  $x^\dagger = A^\dagger y$  denotes the minimum-norm solution of (1.1). In this context, we consider the noise model  $\|y^\delta - y\| \leq \delta$  with noise level  $\delta > 0$ .

The simplest iterative regularization approach for solving (1.1) seems to be the Landweber method, which is given by the iteration procedure

$$x_{k+1}^\delta = x_k^\delta + \Delta t A^*(y^\delta - Ax_k^\delta), \quad \Delta t \in (0, 2/\|A\|^2) \quad (k = 0, 1, 2, \dots) \quad (1.2)$$

with some starting element  $x_0 \in \mathcal{X}$ , where  $A^*$  denotes the adjoint operator of  $A$ . The continuous analog to (1.2) as  $\Delta t$  tends to zero is known as *asymptotic regularization* or (see, e.g., [26, 27]). It is written as a first order evolution equation of the form

$$\dot{x}^\delta(t) + A^*Ax^\delta(t) = A^*y^\delta \quad (1.3)$$

with some initial condition, where an artificial scalar time  $t$  is introduced. There must be chosen an appropriate finite *stopping time*  $T_* = T_*(\delta)$  (a priori choice) or  $T_* = T_*(\delta, y^\delta)$  (a posteriori choice) in order to ensure the regularizing property  $x^\delta(T_*) \rightarrow x^\dagger$  as  $\delta \rightarrow 0$ .

For the asymptotic regularization it is well-known that for all  $p > 0$  under Hölder-type source conditions

$$x^\dagger = (A^*A)^p v, \quad \|v\| \leq \rho, \quad (1.4)$$

and for the stopping time  $T_*$  selected according to the a priori choice  $T_* = T_*(\delta) \sim \delta^{-\frac{2}{2p+1}}$  we have for all  $p > 0$  (cf., e.g., [26, Theorem 2]) the order optimal convergence rate

$$\|x^\delta(T_*) - x^\dagger\| = \mathcal{O}(\delta^{\frac{2p}{2p+1}}) \quad \text{as } \delta \rightarrow 0. \quad (1.5)$$

From the analog result of the Landweber iteration in [5, Theorem 6.5] it can be concluded for asymptotical regularization that we have for the stopping time  $T_* = T_*(\delta, y^\delta)$  chosen according to Morozov's discrepancy principle the formulas

$$\|x^\delta(T_*) - x^\dagger\| = \mathcal{O}(\delta^{\frac{2p}{2p+1}}) \quad \text{and} \quad T_* = \mathcal{O}(\delta^{-\frac{2}{2p+1}}) \quad \text{as } \delta \rightarrow 0. \quad (1.6)$$

Moreover, it has been shown that by using Runge-Kutta integrators, all of the properties of asymptotic regularization (1.3) carry over to its numerical realization [24]. Hence, the continuous model (1.3) is of particular importance for studying the intrinsic properties of a broad class of general linear regularization methods for inverse problems, and can be used for the development of new iterative regularization algorithms by combining some appropriate numerical schemes.

A fatal defect for large-scale problems is the slow performance of Landweber iteration (too many iterations required for optimal stopping) as well as of the asymptotical regularization method, i.e. too excessive stopping times  $T_*$  are required for obtaining optimal convergence rates (1.5). Therefore, in practice, accelerating strategies are usually used. The well-known methods are semi-iterative methods (e.g. the Brakhage's  $\nu$ -method) and the Nesterov acceleration scheme. It has been proven that

- for the  $\nu$ -method, the optimal convergence rates can be obtained with approximately the square root of iterations than those needed for ordinary Landweber iteration [5, § 6.3]. However, in contrast to the Landweber iteration, the  $\nu$ -methods show a saturation phenomenon; i.e., the optimal convergence rate (1.5) and the asymptotic  $k^* = \mathcal{O}(\delta^{-\frac{1}{2p+1}})$  hold only for  $p \leq \nu$  and  $p \leq \nu - 0.5$ , respectively.
- For the Nesterov acceleration scheme, the optimal convergence rates are obtained if  $p \leq 1/2$  and if the iteration is terminated according to an a priori stopping rule. If  $p > 1/2$  or if the iteration is terminated according to the Morozov's conventional discrepancy principle, only sub-optimal convergence rates can be guaranteed.

In both cases, sub-optimal convergence rates can be obtained with the same acceleration speed as by the  $\nu$ -method [22].

Recently, inspired by these two accelerated regularization methods and the advantage of the continuous model, the authors extended (1.3) in [28] to the damped second order dynamics of the form

$$\ddot{x}^\delta(t) + \eta \dot{x}^\delta(t) + A^* A x^\delta(t) = A^* y^\delta, \quad \eta > 0. \quad (1.7)$$

It has been shown that, under condition (1.4), (1.7) exhibits the same convergence rate (1.5) as the asymptotic regularization method for all smoothness index  $p > 0$ . By using the total energy discrepancy principle for choosing the terminating time and the damped symplectic integrators, a new iterative regularization algorithm has been proposed in this regard. Even though numerical experiments demonstrate the acceleration effect of the method (1.7), rigorous mathematical proof is still lacking.

In this paper, we replace the first derivative in the dynamical model (1.3) with appropriate fractional derivatives. Specifically, we consider in Hilbert spaces the following initial value problem

$$\left({}^C D_{0+}^\theta x^\delta\right)(t) + A^* A x^\delta(t) = A^* y^\delta, \quad D^k x^\delta(0) = b_k, \quad k = 0, \dots, n-1, \quad (1.8)$$

to an evolution equation of fractional order  $\theta$ , where  $\theta \in (0, 2)$ ,  $n = \lfloor \theta \rfloor + 1$ , and  $\lfloor \cdot \rfloor$  is the floor function.  $D^k$  denotes the usual differential operator of order  $k$ . The left-side Caputo fractional derivative is defined by  $({}^C D_{0+}^\theta x)(t) := I_{0+}^{n-\theta} D^n x(t)$ , where  $I_{0+}^{n-\theta}$  is the left-side Riemann-Liouville integral operator, i.e.,  $(I_{0+}^{n-\theta} x)(t) := \frac{1}{\Gamma(\theta)} \int_0^t \frac{x(\tau)}{(t-\tau)^{1-n+\theta}} d\tau$  with the gamma function  $\Gamma(\cdot)$ . Note that, for  $\theta = 1$ , (1.8) coincides with Showalter's method (asymptotical regularization). We will call the fractional dynamics (1.8) with an appropriate choice of terminating time *Fractional Asymptotical Regularization (FAR)*. The main goal of this paper is to show that under smoothness assumptions imposed on the exact solution, FAR with  $\theta \in (1, 2)$  yields an accelerated optimal regularization method.

The paper is structured as follows: in Section 2, we review basic concepts in the general linear regularization theory so that they can be applied to the convergence analysis of FAR. The regularization properties of FAR under Hölder-type and logarithmic source conditions are studied in Sections 3 and 4, respectively. Section 5 is devoted to the numerical realization of FAR. Finally, concluding remarks are given in Section 6.

## 2. Linear regularization methods revisited

We start with some definitions, taken from [21] and [12], respectively.

DEFINITION 2.1. A function  $\varphi : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  is called an index function if it is continuous, strictly increasing and satisfies  $\lim_{\lambda \rightarrow 0^+} \varphi(\lambda) = 0$ .

DEFINITION 2.2. A family of functions  $g_\alpha(\lambda)$  ( $0 < \lambda \leq \|A\|^2$ ), defined for regularization parameters  $0 < \alpha \leq \bar{\alpha}$ , is called a generator function for a linear regularization method to the ill-posed linear operator equation (1.1) if the following three conditions are fulfilled:

- (i) For the bias function  $r_\alpha(\lambda) = 1 - \lambda g_\alpha(\lambda)$  we have for any fixed  $\lambda \in (0, \|A\|^2]$  the limit condition  $\lim_{\alpha \rightarrow 0} |r_\alpha(\lambda)| = 0$ .
- (ii) There exists a constant  $\gamma_1 > 0$  such that  $|r_\alpha(\lambda)| \leq \gamma_1$  for all  $\lambda \in (0, \|A\|^2]$  and for all  $\alpha \in (0, \bar{\alpha}]$ .
- (iii) There exists a constant  $\gamma_* > 0$  such that  $\sqrt{\lambda}|g_\alpha(\lambda)| \leq \gamma_*/\sqrt{\alpha}$  for all  $\alpha \in (0, \bar{\alpha}]$ .

Once a generator function family  $g_\alpha(\lambda)$  is chosen, the approximate solutions to (1.1) based on the noisy data  $y^\delta$  are calculated by the procedure

$$x_\alpha^\delta = g_\alpha(A^*A)A^*y^\delta, \quad (2.1)$$

where

$$x_\alpha = g_\alpha(A^*A)A^*Ax^\dagger, \quad (2.2)$$

characterizes the noise-free analog to  $x_\alpha^\delta$ . By condition (iii) of Definition 2.2 and exploiting the triangle inequality, we obtain the well-known error estimates

$$\begin{aligned} \|x_\alpha^\delta - x^\dagger\| &\leq \|x_\alpha - x^\dagger\| + \|x_\alpha^\delta - x_\alpha\| \\ &= \|r_\alpha(A^*A)x^\dagger\| + \delta\|g_\alpha(A^*A)A^*\| \\ &\leq \|r_\alpha(A^*A)x^\dagger\| + \gamma_*\delta/\sqrt{\alpha}. \end{aligned} \quad (2.3)$$

Furthermore, from properties (i) and (ii) of Definition 2.2 we deduce for  $\alpha \rightarrow 0$  point-wise convergence  $r_\alpha(A^*A)x \rightarrow 0$  for any  $x \in \mathcal{X}$  (see, e.g., [5, Theorem 4.1]). If the regularization parameters  $\alpha = \alpha(\delta)$  or  $\alpha = \alpha(\delta, y^\delta)$  are chosen such that

$$\lim_{\delta \rightarrow 0} \alpha = \lim_{\delta \rightarrow 0} \delta^2/\alpha = 0,$$

one can derive convergence  $\|x_\alpha^\delta - x^\dagger\| \rightarrow 0$  as  $\delta \rightarrow 0$  by the estimate (2.3).

If we have an index function  $f$  depending on  $x^\dagger$  as upper bound of the form

$$\|r_\alpha(A^*A)x^\dagger\| \leq f(\alpha) \quad (0 < \alpha \leq \bar{\alpha}) \quad (2.4)$$

then this function also determines the error profile for the solution  $x^\dagger$  in the noisy case, and was therefore termed ‘*profile function*’ in [12]. By using the auxiliary index function  $\Theta(\alpha) := \sqrt{\alpha}f(\alpha)$  and choosing the regularization parameter a priori as  $\alpha_* = \Theta^{-1}(\delta)$ , we can derive under additional

conditions the convergence rate

$$\|x_{\alpha_*}^\delta - x^\dagger\| \leq (1 + \gamma_*)f(\Theta^{-1}(\delta)) = \mathcal{O}(f(\Theta^{-1}(\delta))) \quad \text{as } \delta \rightarrow 0. \quad (2.5)$$

Note that the three requirements in Definition 2.2 are not sufficient to yield profile functions  $f$  for classes of solutions  $x^\dagger$ . It is well-known that convergence rates for approximate solutions of form (2.1) are connected with smoothness conditions imposed on the solution  $x^\dagger$  with respect to the forward operator  $A$ . In order to measure the sensitivity of a regularization method with respect to possible smoothness assumptions, the concept of qualification in the sense of index functions is introduced:

**DEFINITION 2.3.** Let  $\varphi$  be an index function. A regularization method (2.1) generated by  $g_\alpha(\lambda)$  is said to have the qualification  $\varphi$  if there are a constant  $\gamma > 0$ , independent of  $\alpha$ , and a value  $\bar{\alpha} > 0$  such that for all  $\alpha \in (0, \bar{\alpha}]$  the inequality

$$\sup_{\lambda \in (0, \|A\|^2]} |r_\alpha(\lambda)|\varphi(\lambda) \leq \gamma\varphi(\alpha)$$

is satisfied.

It should be noted that the concept of Definition 2.3 is an alternative to the traditional concept of qualification in the sense of a positive number or infinity used in [27] and [5, Chap. 4]. This alternative concept was originally introduced in [21], but it has been frequently used by other authors recently. Only if the index function  $\varphi$  is a qualification in the sense of Definition 2.3 for the regularization method with generator functions  $g_\alpha$ , we can derive the rate result (2.5) from (2.4).

**EXAMPLE 2.1** (Hölder source conditions). In this example we consider the index functions  $\varphi(\lambda) = \lambda^p$  associated with the Hölder source conditions (1.4). If such  $\varphi$  is for some  $p > 0$  a qualification for method generated by  $g_\alpha$ , then we obtain from (2.3)

$$\|x_\alpha^\delta - x^\dagger\| \leq \gamma\alpha^p\rho + \gamma_*\delta/\sqrt{\alpha}$$

and hence for  $\alpha_* = \alpha(\delta) = c\delta^{2p/(2p+1)}$  ( $c > 0$ ) the error estimate and convergence rate

$$\|x_{\alpha_*}^\delta - x^\dagger\| \leq (\gamma c^p + \gamma_*/\sqrt{c})\delta^{\frac{2p}{2p+1}} = \mathcal{O}\left(\delta^{\frac{2p}{2p+1}}\right) \quad \text{as } \delta \rightarrow 0. \quad (2.6)$$

If a benchmark source condition  $\psi$  is known to be a qualification for the method generated by  $g_\alpha$ , then other index functions  $\varphi$  are also qualifications whenever they are *covered* by  $\psi$ , and we refer to [21, Def. 2]

and [21, Prop. 3, Remark 5 and Lemma 2] for the following definition and proposition, respectively.

**DEFINITION 2.4.** Let  $\psi(\lambda)$  ( $0 < \lambda \leq \|A\|^2$ ) be an index function. Then an index function  $\varphi(\lambda)$  ( $0 < \lambda \leq \|A\|^2$ ) is said to be covered by  $\psi$  if there is  $\underline{c} > 0$  such that

$$\underline{c} \frac{\psi(\alpha)}{\varphi(\alpha)} \leq \inf_{\alpha \leq \lambda \leq \|A\|^2} \frac{\psi(\lambda)}{\varphi(\lambda)} \quad (0 < \alpha \leq \bar{\alpha}).$$

**PROPOSITION 1.** The index function  $\varphi$  is a qualification for the method generated by  $g_\alpha$  if  $\varphi$  is covered by  $\psi$  and if  $\psi$  is a qualification for that method. If the quotient function  $\lambda \mapsto \frac{\psi(\lambda)}{\varphi(\lambda)}$  is increasing for  $0 < \lambda \leq \bar{\lambda}$  and some  $\bar{\lambda} > 0$ , then  $\varphi$  is covered by  $\psi$ . If, in particular, the index function  $\varphi(\lambda)$  is concave for  $0 < \lambda \leq \bar{\lambda}$ , then  $\varphi$  is covered by  $\psi(\lambda) = \lambda$ .

**EXAMPLE 2.2 (Logarithmic source conditions).** The focus of this example is on logarithmic source conditions

$$x^\dagger = \varphi_\mu(A^*A)v, \quad \|v\| \leq \rho, \quad (2.7)$$

where for exponents  $\mu > 0$  the function  $\varphi_\mu$  is defined as

$$\varphi_\mu(\lambda) = \begin{cases} \log^{-\mu}(1/\lambda) & \text{for } 0 < \lambda \leq e^{-\mu-1}, \\ \text{arbitrarily extended as index function} & \text{for } \lambda > e^{-\mu-1}. \end{cases} \quad (2.8)$$

Note that the condition (2.7) for any  $\mu > 0$  is significantly weaker than the condition (1.4), even for arbitrarily small  $p > 0$ . As a consequence, the expected logarithmic convergence rates are also significantly lower than in the Hölder case (2.6) and typically occur for severely ill-posed problems (cf., e.g., [13]). If  $\varphi_\mu$  is for some  $\mu > 0$  a qualification for method generated by  $g_\alpha$ , then we obtain from (2.3)

$$\|x_\alpha^\delta - x^\dagger\| \leq \gamma \varphi_\mu(\alpha) \rho + \gamma_* \delta / \sqrt{\alpha},$$

for sufficiently small  $\alpha > 0$ , and hence for  $\alpha_* = \alpha(\delta) = c\delta^\kappa$  ( $c > 0$ ,  $0 < \kappa < 2$ ) the convergence rate

$$\|x_{\alpha_*}^\delta - x^\dagger\| = \mathcal{O}(\varphi_\mu(\delta)) \quad \text{as } \delta \rightarrow 0 \quad (2.9)$$

as a consequence of  $\lim_{\delta \rightarrow 0} \delta^\eta / \varphi_\mu(\delta) = 0$ , which is valid for all  $\eta > 0$ .

### 3. FAR under Hölder-type source conditions

In this section, we show that the FAR method with fractional order  $\theta > 0$  introduced in the introduction can be assigned to the general linear regularization schema recalled in Section 2. Therefore, we start with the

verification of the generator function  $g_\alpha$  occurring in formula (2.1) associated with the fractional differential equation (1.8) and the corresponding regularization properties. For simplicity, let  $b_k = 0$  in (1.8) for all  $k$  under consideration. The case with non-vanishing initial data can be analyzed similarly, see [28].

Let  $\{\sigma_j; u_j, v_j\}_{j=1}^\infty$  be the well-defined singular system for the compact linear operator  $A$ , i.e. we have  $Au_j = \sigma_j v_j$  and  $A^*v_j = \sigma_j u_j$  with ordered singular values  $\|A\| = \sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_j \geq \sigma_{j+1} \geq \dots \rightarrow 0$  as  $j \rightarrow \infty$ . Since the eigenelements  $\{u_j\}_{j=1}^\infty$  and  $\{v_j\}_{j=1}^\infty$  form complete orthonormal systems (with the exception of null-spaces) in  $\mathcal{X}$  and  $\mathcal{Y}$ , respectively, (1.7) is equivalent to

$$\left\langle \left( {}^C D_{0+}^\theta x^\delta \right) (t), u_j \right\rangle + \sigma_j^2 \langle x^\delta(t), u_j \rangle = \sigma_j \langle y^\delta, v_j \rangle, \quad j = 1, 2, \dots$$

Using the decomposition  $x^\delta(t) = \sum_{j=1}^\infty \xi_j(t) u_j$ , we obtain

$$\left( {}^C D_{0+}^\theta \xi_j \right) (t) + \sigma_j^2 \xi_j(t) = \sigma_j \langle y^\delta, v_j \rangle, \quad j = 1, 2, \dots \quad (3.1)$$

On the other hand, according to [15, § 4.1.3], the solution to the equation  $\left( {}^C D_{0+}^\theta \eta \right) (t) - \lambda \eta(t) = \kappa(t)$  with the vanishing initial data is explicitly given by (see formula (4.1.62) in [15])

$$\eta(t) = \int_0^t (t - \tau)^{\theta-1} E_{\theta, \theta}(\lambda(t - \tau)^\theta) \kappa(\tau) d\tau, \quad (3.2)$$

where the two-parametric Mittag-Leffler function  $E_{\theta_1, \theta_2}(z)$  is defined as

$$E_{\theta_1, \theta_2}(z) = \sum_{k=0}^\infty \frac{z^k}{\Gamma(\theta_1 k + \theta_2)}.$$

Therefore, together with (3.2), we deduce that the solution of (3.1) with the vanishing initial data is given by

$$\xi_j(t) = \sigma_j \langle y^\delta, v_j \rangle \int_0^t (t - \tau)^{\theta-1} E_{\theta, \theta}(-\sigma_j^2(t - \tau)^\theta) d\tau,$$

Consequently, by the above result, together with the identity [8, § 4.4]

$$\int_0^t E_{\theta_1, \theta_2}(\xi z^{\theta_1}) z^{\theta_2-1} dz = t^{\theta_2} E_{\theta_1, \theta_2+1}(\xi t^{\theta_1}), \quad \theta_2 > 0,$$

we deduce for all  $\theta > 0$  that

$$\xi_j(t) = \sigma_j t^\theta E_{\theta, \theta+1}(-\sigma_j^2 t^\theta) \langle y^\delta, v_j \rangle, \quad j = 1, 2, \dots$$



Thus, by using  $x^\delta(t) = \sum_{j=1}^{\infty} \xi_j(t)u_j$ , we obtain the explicit formula

$$x^\delta(t) = t^\theta \sum_{j=1}^{\infty} E_{\theta, \theta+1}(-\sigma_j^2 t^\theta) \sigma_j \langle y^\delta, v_j \rangle u_j = g^\theta(t, A^*A) A^* y^\delta \quad (3.3)$$

for the solution to (1.7), where the generator function  $g^\theta(t, \lambda)$  characterizing the FAR method attains the form

$$g^\theta(t, \lambda) := t^\theta E_{\theta, \theta+1}(-\lambda t^\theta). \quad (3.4)$$

Furthermore, by using the recurrence relations

$$z^m E_{\theta_1, \theta_2 + m\theta_1}(z) = E_{\theta_1, \theta_2}(z) - \sum_{k=0}^{m-1} \frac{z^k}{\Gamma(\theta_1 k + \theta_2)}$$

we obtain an associated bias function  $r^\theta(t, \lambda) = 1 - \lambda g^\theta(t, \lambda)$  of the form

$$r^\theta(t, \lambda) = 1 - \lambda t^\theta E_{\theta, \theta+1}(-\lambda t^\theta) = E_{\theta, 1}(-\lambda t^\theta) \equiv E_\theta(-\lambda t^\theta), \quad (3.5)$$

where  $E_\theta(z) = \sum_{k=0}^{\infty} \frac{z^k}{\Gamma(\theta k + 1)}$  denotes the classical Mittag-Leffler Function.

**REMARK 3.1.** For  $\theta = 0.5$ , we have  $r(t, \lambda) = e^{-\lambda^2 t} [1 + \operatorname{erf}(-\lambda\sqrt{t})]$ , where  $\operatorname{erf}(z) := \frac{2}{\sqrt{\pi}} \int_0^z e^{-\tau^2} d\tau$  represents the error function. Moreover,  $r(t, \lambda) = e^{-\lambda t}$  for  $\theta = 1$ , while  $r(t, \lambda) = \cos(\sqrt{\lambda t})$  when  $\theta = 2$ .

**LEMMA 3.1.** *Let  $\theta \in (0, 2)$  be a fixed number. Then we have for all  $z \in \mathbb{R}_+$  the following three inequalities*

$$|E_{\theta, \theta+1}(-z)| \leq C_\theta / (1 + z), \quad (3.6)$$

$$|E_\theta(-z)| \leq C_\theta / (1 + z), \quad (3.7)$$

$$|E_\theta(-z)| \leq 3, \quad (3.8)$$

where  $C_\theta > 0$  is a constant, depending only on  $\theta$ .

**P r o o f.** Inequalities (3.6) and (3.7) follow from [23, Theorem 1.6] and [8, Corollary 3.7] respectively. For the estimate (3.8), we distinguish three cases: (I)  $\theta \in (0, 1)$ , (II)  $\theta = 1$ , and (III)  $\theta \in (1, 2)$ . In this first case, (3.8) holds according to the uniform estimate [25, Theorem 4]:

$$\frac{1}{1 + \Gamma(1 - \theta)z} \leq E_\theta(-z) \leq \frac{1}{1 + \Gamma(1 + \theta)^{-1}z}, \quad (3.9)$$

For the second case, we have  $E_1(-z) = e^{-z} \leq 1$ . Now, we consider the last case. To this end, we set  $z := t^\theta$  and consider the function  $E_\theta(-t^\theta)$ .

It is well known (see, e.g., [20]) that  $E_\theta(-t^\theta) = f_\theta(t) + h_\theta(t)$ , where the completely monotone function  $-f_\theta$  is defined by

$$f_\theta(t) = \frac{1}{\pi} \int_0^\infty e^{-rt} \frac{r^{\theta-1} \sin(\theta\pi)}{r^{2\theta} + 2r^\theta \cos(\theta\pi) + 1} dr, \quad (3.10)$$

and the oscillatory part  $h_\theta$  is given by

$$h_\theta(t) = \frac{2}{\theta} e^{t \cos(\pi/\theta)} \cos(t \sin(\pi/\theta)).$$

Obviously, for  $\theta \in (1, 2)$ ,  $f_\theta$  is a monotonically increasing function such that  $-1 < 1 - \frac{2}{\theta} = \frac{1}{\pi} \int_0^\infty \frac{r^{\theta-1} \sin(\theta\pi)}{r^{2\theta} + 2r^\theta \cos(\theta\pi) + 1} dr \leq f_\theta(t) \leq 0$ , while  $|h_\theta(t)| \leq 2/\theta$  as  $\cos(\pi/\theta) < 0$ . This implies  $|E_\theta(-t^\theta)| \leq 3$ .  $\square$

We remark that  $C_\theta \rightarrow 1$  as  $\theta \searrow 1$ , and  $C_\theta \rightarrow \infty$  as  $\theta \nearrow 2$ .

In order to assign FAR to the general linear regularization schema introduced in Section 2, we exploit the one-to-one correspondence between the artificial time  $t > 0$  and the conventional regularization parameter  $\alpha > 0$  by setting  $t := \alpha^{-1/\theta}$ .

**THEOREM 3.1.** *For  $\theta \in (0, 2)$ ,  $g_\alpha(\lambda) := g^\theta(\alpha^{-1/\theta}, \lambda)$  represents a generator function in the sense of Definition 2.2 with  $\theta_* = C_\theta/2$  and  $C_\theta$  from Lemma 3.1. Consequently, FAR is a linear regularization method for the ill-posed linear operator equation (1.1).*

**P r o o f.** We have to check the three requirements (i), (ii) and (iii) in Definition 2.2 for all  $\theta \in (0, 2)$ .

Since  $r_\alpha(\lambda) = r^\theta(\alpha^{-1/\theta}, \lambda) = E_\theta(-\lambda\alpha^{-1})$ , it follows from (3.7) that

$$|r_\alpha(\lambda)| \leq C_\theta \alpha / (\alpha + \lambda), \quad (3.11)$$

which shows that (i) is satisfied. The second condition (ii) is a consequence of (3.8). It remains to find a bound  $\gamma_*$  in (iii). Note that  $g_\alpha(\lambda) = \alpha^{-1} E_{\theta, \theta+1}(-\lambda\alpha^{-1})$ . By using (3.6), we obtain the inequalities

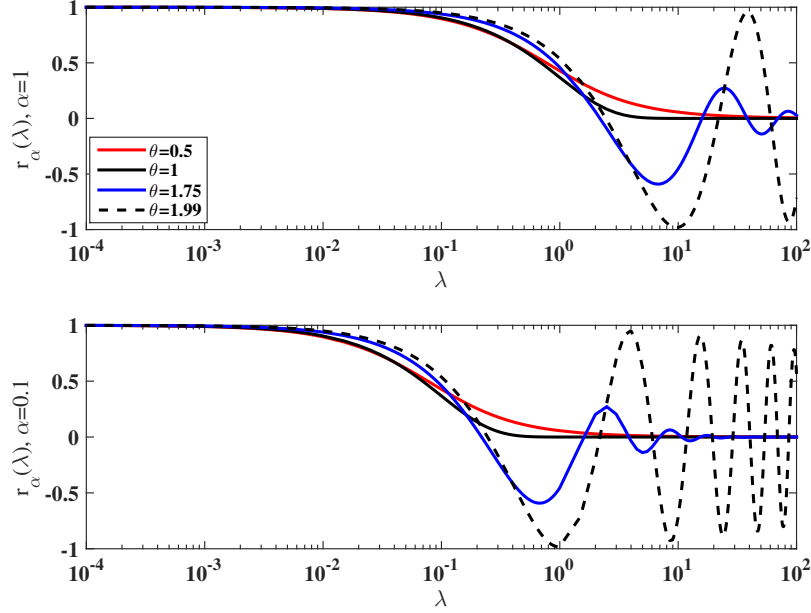
$$\sqrt{\lambda} |g_\alpha(\lambda)| \leq \sqrt{\lambda} C_\theta / (\alpha + \lambda) \leq C_\theta / (2\sqrt{\alpha}) \quad (3.12)$$

and hence  $\gamma_* = C_\theta/2$ , which completes the proof.  $\square$

Referring back to the time variable  $t$  we obtain a linear regularization method for the ill-posed operator equation (1.1) with the procedure

$$x^\delta(t) = g_{t^{-\theta}}(A^* A) A^* y^\delta. \quad (3.13)$$

The behavior of the bias function  $r_\alpha(\lambda)$ , or equivalently  $r_{t^{-\theta}}(\lambda)$ , is visualized in Figure 1. It follows from Remark 3.1 that for  $\theta = 2$ ,  $r_{t^{-2}}(\lambda) = \cos(\sqrt{\lambda t^2})$ . Obviously, we have for all  $\lambda$  under consideration  $r_{t^{-2}}(\lambda) \not\rightarrow 0$  as  $t \rightarrow \infty$ , which damages the first condition of Definition 2.2. In order to

FIGURE 1. Behavior of bias function  $r_\alpha(\lambda)$  for varying  $\theta$ .


obtain a regularization method in the case  $\theta = 2$ , an additional damping term in the model (1.8) should be introduced, see [28] for details.

**PROPOSITION 2.** For all  $\theta \in (0, 1) \cup (1, 2)$  the function  $\varphi(\lambda) = \lambda^p$  is a qualification of the FAR method if and only if  $0 < p \leq 1$ , i.e. there is a saturation for  $p = 1$ .

**P r o o f.** The qualification of the FAR can be deduced from the estimate (3.7) and the following inequalities

$$r_\alpha(\lambda) = |E_\theta(-\lambda\alpha^{-1})\lambda^p| \leq \frac{C_\theta\lambda^p}{1 + \lambda\alpha^{-1}} \leq \gamma(p, \theta)\alpha^{\min(p, 1)}, \quad (3.14)$$

where  $\gamma(p, \theta) = C_\theta p^p(1-p)^{1-p}$  for  $0 < p < 1$ , and  $\gamma(p, \theta) = C_\theta \|A\|^{2(p-1)}$  for  $1 \leq p < 2$ . The saturation at  $p = 1$  follows from the lower bound in (3.9) and the asymptotical behavior, see e.g. [23, § 1.2],  $E_\theta(-z)z \rightarrow 1/\Gamma(1 - \theta)$  as  $z \rightarrow \infty$ .  $\square$

REMARK 3.2. By the well-known results of Showalter's method, see e.g. [28, Remark 2.2],  $\varphi(\lambda) = \lambda^p$  is a qualification of the FAR method with parameter  $\theta = 1$  for all  $p > 0$  without saturation.

THEOREM 3.2. *Let  $x(t)$  and  $x^\delta(t)$  be solutions of (1.8) with noise-free and noisy data, respectively. Then, under the assumption  $x^\dagger \in \mathcal{R}((A^*A)^p)$  (cf. (1.4)), we have for the FAR method in the case  $\theta \in (0, 1) \cup (1, 2)$  and  $p \in (0, 1]$  the convergence rate*

$$\|x^\delta(T_*) - x^\dagger\| = \mathcal{O}\left(\delta^{\frac{2p}{2p+1}}\right) \quad \text{as } \delta \rightarrow 0 \quad (3.15)$$

whenever the terminating time  $T_* = T_*(\delta)$  is chosen according to

$$T_*(\delta) \sim \delta^{-\frac{2}{\theta(2p+1)}}. \quad (3.16)$$

Moreover, the following additional assertions are true if  $\theta \in (0, 1)$ :

- (i) *The convergence rate of method error  $\|x(t) - x^\dagger\| = \mathcal{O}(t^{-\theta p})$  holds, if and only if  $x^\dagger$  satisfies a source condition of type (1.4).*
- (ii) *The worst-case convergence rate*

$$\sup_{y^\delta: \|y^\delta - y\| \leq \delta} \inf_{t \in \mathbb{R}_+} \|x^\delta(t; y^\delta) - x^\dagger\|^2 = \mathcal{O}\left(\delta^{\frac{2p}{2p+1}}\right) \quad (3.17)$$

is equivalent to the following weaker condition

$$\|E_s x^\dagger\|^2 := \int_0^s 1d\|E_\lambda x^\dagger\|^2 = \mathcal{O}(s^{2p}). \quad (3.18)$$

- (iii) *If a terminating time  $T_*$  exists such that*

$$\sup_{y^\delta: \|y^\delta - y\| \leq \delta} \|x^\delta(T_*; y^\delta) - x^\dagger\|^2 = o\left(\delta^{\frac{2p}{2p+1}}\right), \quad (3.19)$$

then  $x^\dagger = 0$ .

PROOF. The assertions concerning the convergence rate (3.15) follow directly from (2.6) by setting  $T_*(\delta) = (\alpha_*)^{-1/\theta}$  in combination with Proposition 2. The remaining assertions (i), (ii) and (iii) of the theorem are based on standard arguments of regularization theory ([5, § 4.2] and [6]) in light of estimates (3.6), (3.9), and (3.11).  $\square$

REMARK 3.3. (a) The first part of Theorem 3.2 means that the optimal convergence rates (3.15) of FAR can be obtained with approximately  $1/\theta$  of iterations than needed for ordinary asymptotical regularization method [5]. Therefore, FAR with  $\theta \in (1, 2)$  yields an accelerated regularization method for all  $p > 0$ . As the estimate (2.6) shows, for fixed  $\delta > 0$  the error

of regularization is proportional to  $\gamma$  and  $\gamma_*$ . These constants, however, tend to infinity as  $\theta \nearrow 2$  such that values  $\theta$  close to 2 need not be the best choices of this parameter. (b) The source condition (1.4) implies the weaker source condition (3.18). The converse is in general not true, however, (3.18) implies  $x^\dagger \in \cap_{\nu < p} \mathcal{R}((A^*A)^\nu)$ , see [5, Lemma 4.12].

In practice, the a priori stopping rule (3.16) in Theorem 3.2 is not realistic, since a good terminating time  $T_*$  requires in addition to the knowledge of  $\delta$  also the knowledge of  $p$  characterizing the specific smoothness of the unknown exact solution  $x^\dagger$ . Therefore, a posteriori parameter choices of the stopping parameter are preferred, and we consider Morozov's discrepancy principle as the most prominent version exploiting zeros of the discrepancy function

$$\chi(t) = \|Ax^\delta(t) - y^\delta\| - \tau\delta, \quad (3.20)$$

where we assume  $\tau > 3$  for the occurring factor of the noise level  $\delta$ .

**LEMMA 3.2.** *If for  $\tau > 3$  the condition  $\|y^\delta\| > \tau\delta$  is satisfied, the function  $\chi(T)$  has at least one root.*

**P r o o f.** By the explicit formula of  $x^\delta(t)$ , c.f. (3.13), we derive together with the estimates (3.8) and (3.14) that

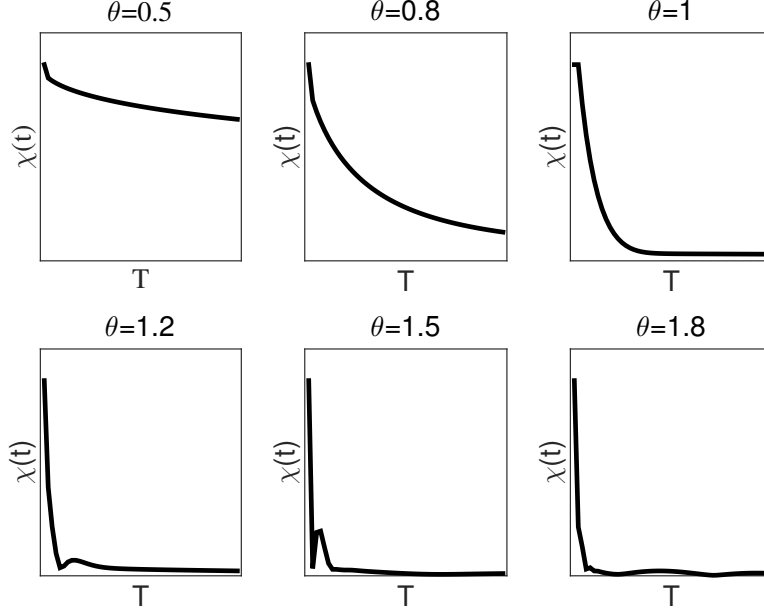
$$\begin{aligned} \chi(t) &= \|Ax^\delta(t) - y^\delta\| - \tau\delta = \|r_{t-\theta}(AA^*)(y - (y - y^\delta))\| - \tau\delta \\ &\leq \|r_{t-\theta}(AA^*)Ax^\dagger\| + \|r_{t-\theta}(AA^*)(y - y^\delta)\| - \tau\delta \\ &\leq \|x^\dagger\| \sup_{\lambda \in (0, \|A\|^2]} |r_{t-\theta}(\lambda)|\sqrt{\lambda} - (\tau - 3)\delta \\ &\leq \gamma(p, \theta)\|x^\dagger\|t^{-\theta/2} - (\tau - 3)\delta \rightarrow -(\tau - 3)\delta < 0 \end{aligned}$$

as  $t \rightarrow \infty$ . The continuity of  $\chi(t)$  is obvious as we are dealing with the linear problem. Since  $\chi(0) = \|y^\delta\| - \tau\delta > 0$ , the existence of the root of  $\chi(t)$  follows from the Bolzano's theorem.  $\square$

**REMARK 3.4.** As the proof of Lemma 3.2 indicates, the function  $\chi(t)$  is bounded above by a decreasing function  $\zeta(t) := \gamma(p, \theta)\|x^\dagger\|t^{-\theta/2} - (\tau - 3)\delta$ . Thus, the trend of  $\chi(t)$  is to be a decreasing function, where oscillations may occur, and we refer to Figure 2 for illustration.

**THEOREM 3.3.** *Under the assumption  $x^\dagger \in \mathcal{R}((A^*A)^p)$ , we have for the terminating time  $T_* = T_*(\delta, y^\delta)$  of FAR chosen as a solution of  $\chi(T) = 0$*

FIGURE 2. The behavior of  $\chi(T)$  from (3.20) with different fractional orders  $\theta$ .



(cf. 3.20) the convergence rates

$$T_* = \mathcal{O}\left(\delta^{-\frac{2}{\theta(2p+1)}}\right) \quad \text{and} \quad \|x^\delta(T_*) - x^\dagger\| = \mathcal{O}\left(\delta^{\frac{2p}{2p+1}}\right) \quad (3.21)$$

for  $p \in (0, 1]$  if  $\theta \in (0, 1)$  and for all  $p > 0$  if  $\theta \in [1, 2)$ .

*P r o o f.* Using the interpolation inequality  $\|B^p u\| \leq \|B^q u\|^{p/q} \|u\|^{1-p/q}$  and the source conditions  $x^\dagger = (A^* A)^p v$ , we deduce that

$$\begin{aligned} \|r_\alpha(A^* A)x^\dagger\| &= \|(A^* A)^p r_\alpha(A^* A)v\| \\ &\leq \|(A^* A)^{(p+1/2)} r_\alpha(A^* A)v\|^{2p/(2p+1)} \cdot \|r_\alpha(A^* A)v\|^{1/(2p+1)} \\ &= \|Ar_\alpha(A^* A)x^\dagger\|^{2p/(2p+1)} \cdot \|r_\alpha(A^* A)v\|^{1/(2p+1)}. \end{aligned} \quad (3.22)$$

Since  $T_*$  is chosen according to the equation  $\chi(T) = 0$ , we derive that

$$\begin{aligned} \tau\delta &= \|Ax^\delta(T_*) - y^\delta\| \\ &= \left\| Ar_{T_*-\theta}(A^* A)x^\dagger + r_{T_*-\theta}(AA^*)(y^\delta - y) \right\| \\ &\geq \left| \|Ar_{T_*-\theta}(A^* A)x^\dagger\| - \|r_{T_*-\theta}(AA^*)(y^\delta - y)\| \right|. \end{aligned} \quad (3.23)$$

Now we combine the estimates (3.22) and (3.23) to obtain, with the source conditions, that

$$\begin{aligned} \|r_\alpha(A^*A^*)x^\dagger\| &\leq \|Ar_\alpha(A^*A)x^\dagger\|^{2p/(2p+1)} \cdot \|r_\alpha(A^*A)v\|^{1/(2p+1)} \\ &\leq \left( \tau\delta + \|r_{T_*^{-\theta}}(AA^*)(y^\delta - y)\| \right)^{2p/(2p+1)} \rho^{1/(2p+1)} \\ &\leq c'_2 \rho^{1/(2p+1)} \delta^{2p/(2p+1)} \end{aligned} \quad (3.24)$$

where  $c'_2 := (\tau + 3)^{2p/(2p+1)}$ .

On the other hand, in a similar fashion to (3.23), it is easy to show that

$$\begin{aligned} \tau\delta &= \|Ax^\delta(T_*) - y^\delta\| = \|r_{T_*^{-\theta}}(AA^*)y + r_{T_*^{-\theta}}(AA^*)(y^\delta - y)\| \\ &\leq \|Ar_{T_*^{-\theta}}(A^*A)x^\dagger\| + \|r_{T_*^{-\theta}}(AA^*)(y^\delta - y)\| \\ &\leq \|Ar_{T_*^{-\theta}}(A^*A)x^\dagger\| + 3\delta. \end{aligned} \quad (3.25)$$

If we combine the above inequality with the source conditions and the qualification inequality, we obtain

$$\begin{aligned} (\tau - 3)\delta &\leq \|Ar_{T_*^{-\theta}}(A^*A)x^\dagger\| \\ &\leq \|(A^*A)^{p+1/2}r_{T_*^{-\theta}}(A^*A)v\| \leq 2\rho\gamma(T_*)^{-\theta(2p+1)/2}, \end{aligned} \quad (3.26)$$

which yields the estimate for  $T_*$  in (3.21). Finally, using (2.3), the estimate for  $T_*$  and (3.24), we conclude that

$$\begin{aligned} \|x^\delta(T_*) - x^\dagger\| &\leq \|r_\alpha(A^*A^*)x^\dagger\| + \gamma_* \sqrt{T_*^\theta} \delta \\ &\leq c_1 \rho^{1/(2p+1)} \delta^{2p/(2p+1)} + \gamma_* \left( \frac{2\gamma}{\tau-3} \right)^{1/(2p+1)} \rho^{1/(2p+1)} \delta^{2p/(2p+1)} \\ &= C_1 \rho^{1/(2p+1)} \delta^{2p/(2p+1)}. \end{aligned}$$

This completes the proof.  $\square$

#### 4. FAR under logarithmic source conditions

Now, we turn to the study of FAR method under conditions (2.7).

**PROPOSITION 3.** For all  $\theta \in (0, 2)$  and all  $\mu > 0$  the index function  $\varphi_\mu(\lambda)$ , defined in (2.8), is a qualification of the FAR method.

**P r o o f.** For all  $\theta \in (0, 2)$  the function  $\psi(\lambda) = \lambda$  is a qualification of the FAR method, as shown in Proposition 2. Moreover, for arbitrary  $\mu > 0$  the index function  $\varphi_\mu(\lambda)$  is concave for all  $0 < \lambda \leq e^{-\mu-1}$ , hence due to Proposition 1 covered by  $\psi$  and consequently also a qualification of the FAR method. This proves the proposition.  $\square$

Based on the above proposition, the following theorem holds as a consequence of the discussions in Example 2.2 (cf. formula (2.9)).

**THEOREM 4.1.** *Let  $x^\delta(t)$  denote the solution of (1.8) with noisy data  $y^\delta$  and noise level  $\delta > 0$ . Then, under the assumption  $x^\dagger \in \mathcal{R}(\varphi_\mu(A^*A))$  (cf. (2.7)) for arbitrary  $\mu > 0$ , we have for the FAR method and all  $0 < \theta < 2$  the convergence rate*

$$\|x^\delta(T_*) - x^\dagger\| = \mathcal{O}(\log^{-\mu}(\delta^{-1})) \quad \text{as } \delta \rightarrow 0 \quad (4.1)$$

whenever the terminating time  $T_*$  of FAR is chosen according to

$$T_*(\delta) \sim \delta^{-\kappa/\theta},$$

for some  $0 < \kappa < 2$ ,

Now, consider the a posteriori choice of regularization parameter. We start with the following lemma.

**LEMMA 4.1.** *The index function  $\sqrt{\lambda}\varphi_\mu(\lambda)$  is a qualification of FAR, where  $\varphi_\mu$  is defined in (2.8). Consequently, a constant  $c_3$  exists such that*

$$\sup_{\lambda \in (0, e^{-2\mu-1})} \sqrt{\lambda} |r_\alpha(\lambda)| \varphi_\mu(\lambda) \leq c_3 \sqrt{\alpha} \log^{-\mu}(\alpha^{-1}).$$

**P r o o f.** The assertion of the lemma follows from Proposition 3 in combination with the fact that  $\varphi(\lambda) = \lambda$  is a qualification of FAR for all  $\theta \in (0, 2)$ , and with the following inequality

$$\sup_{\lambda \in (0, e^{-2\mu-1})} \left( \sqrt{\lambda} r_\alpha(\lambda) \varphi_\mu(\lambda) \right)^2 \leq \sup_{\lambda \in (0, e^{-2\mu-1})} \lambda r_\alpha(\lambda) \sup_{\lambda \in (0, e^{-2\mu-1})} r_\alpha(\lambda) \varphi_{2\mu}(\lambda).$$

□

**THEOREM 4.2.** *Under the assumptions of Theorem 4.1, but for the terminating time  $T_* = T_*(\delta, y^\delta)$  of FAR chosen as a solution of  $\chi(T) = 0$  (cf. 3.20)), we have*

$$T_* = \mathcal{O} \left( W_{2\mu/\theta} \left( \delta^{-2/\theta} \log^{-2/\theta}(\delta^{-1}) \right) \right) \quad \text{as } \delta \rightarrow 0, \quad (4.2)$$

where  $W_\alpha(z)$  is the unique solution of  $\zeta \log^\alpha(\zeta) = z$  with respect to  $\zeta$ , and moreover the convergence rate

$$\|x^\delta(T_*) - x^\dagger\| = \mathcal{O}(\log^{-\mu}(\delta^{-1})) \quad \text{as } \delta \rightarrow 0. \quad (4.3)$$

**P r o o f.** Without loss of generality we assume that  $\delta < e^{-2\mu-1}$ . By combining inequality (3.25) and Lemma 4.1 with  $\alpha$  replaced by  $T_*^{-\theta}$ , we



obtain

$$\begin{aligned} (\tau - 3)\delta &\leq \|Ar_{T_*^{-\theta}}(A^*A)x^\dagger\| \leq c_3 T_*^{-\theta/2} \log^{-\mu}(T_*^\theta) \\ &\leq c_3 \theta^{-\mu/2} T_*^{-\theta/2} \log^{-\mu}(T_*). \end{aligned}$$

Therefore, for  $\delta \in (0, e^{-1})$ , and if we set  $c_4 := c_3 \theta^{-\mu/2} (\tau - 3)^{-1}$ , we have

$$T_* \log^{2\mu/\theta}(T_*) \leq c_4^{2/\theta} \delta^{-2/\theta} \log^{-2/\theta}(\delta^{-1}), \quad (4.4)$$

which yields the estimate (4.2), as well as the inequality

$$\sqrt{T_*^\theta} \leq c_4 \delta^{-1} \log^{-1-\mu}(\delta^{-1}). \quad (4.5)$$

The estimate (4.3) can be driven according to inequalities (2.3), (4.5), and to the Proposition 3.  $\square$

REMARK 4.1. (a) To the best of our knowledge, under the logarithmic source condition, for obtaining the optimal convergence rate (4.3), existing regularization methods with a posteriori regularization parameter selection methods require the time cost  $T_{old} = \mathcal{O}(\delta^{-2} \log^{-2-2\mu}(\delta^{-1}))$ . However, in the case of the FAR method, according to (4.4) we derive for  $\theta > 1$  that

$$T_* \ll T_* \log^{2\mu/\theta}(T_*) \leq c_4^{2/\theta} \delta^{-2/\theta} \log^{-2/\theta}(\delta^{-1}) \ll T_{old}$$

for sufficient small  $\delta > 0$ . This means that FAR with  $1 < \theta < 2$  yields a super accelerated optimal regularization method. (b) Similar to Theorem 3.2, the converse result for FAR under logarithmic source conditions can be established by the existing technique, proposed in e.g. [1].

## 5. Numerical investigation of the FAR method

Roughly speaking, the fractional differential equation (1.8) with appropriate numerical discretization schemes for the artificial time variable and stopping rule of iteration steps yields a concrete iterative regularization method. Rather than performing a rigorous numerical analysis, we provide in this section a numerical example to demonstrate if the regularization property and the acceleration effect of the FAR carry over to its numerical realization. To this end, we consider the one-step Adams-Moulton method [2, 3], namely,

$$\begin{cases} x_{k+1}^{\delta, P} = \frac{1}{\Gamma(\theta)} \sum_{j=0}^k b_{j, k+1} A^*(y^\delta - Ax_j^\delta), \\ x_{k+1}^\delta = \frac{1}{\Gamma(\theta)} \left( a_{k+1, k+1} A^*(y^\delta - Ax_{k+1}^{\delta, P}) + \sum_{j=0}^k a_{j, k+1} A^*(y^\delta - Ax_j^\delta) \right), \end{cases} \quad (5.1)$$

where  $b_{j,k+1} = \frac{\Delta t^\theta}{\theta} [(k-j+1)^\theta - (k-j)^\theta]$ , and  $a_{j,k+1} = \frac{\Delta t^\theta}{\theta(\theta+1)} \cdot d_{j,k+1}$ ,

$$d_{j,k+1} = \begin{cases} [k^{\theta+1} - (k-\theta)(k+1)^\theta], & j = 0, \\ [(k-j+2)^{\theta+1} + (k-j)^{\theta+1} - 2(k-j+1)^{\theta+1}], & 1 \leq j \leq k \\ 1, & j = k+1. \end{cases}$$

REMARK 5.1. As the Landweber iterates, the iterates  $x_{k+1}^\delta$  of (5.1) obviously belong to the Krylov subspace  $\text{Span} \{A^*y^\delta, \dots, (A^*A)^{k-1}A^*y^\delta\}$ . Therefore, the solution  $x_{k+1}^\delta$  of (5.1) can be written as  $x_{k+1}^\delta = g_k(A^*A)A^*y^\delta$ , where  $g_k$  is a polynomial of degree  $k$ . For proving the regularization property of (5.1), one has to check that  $g_k(\lambda)$  fulfills all three conditions in Definition 2.2. Moreover, in order to prove the acceleration effect of 2.2 under source conditions (1.4), one must show that for all  $\lambda \in (0, \|A\|^2]$ :  $|r_k(\lambda)|(k\lambda)^p \rightarrow 0$  as  $k \rightarrow \infty$ , where  $r_k(\lambda) = 1 - \lambda g_k(\lambda)$  represents the residual polynomial. However, all of these issues are not addressed here since it is out of the scope of our current paper. Rather we refer to a similar result in [7] for the second order asymptotical regularization.

Now, by employing the newly developed iterative regularization method (5.1), we present some numerical results for the following integral equation

$$Ax(s) := \int_0^1 K(s,t)x(t)dt = y(s), \quad K(s,t) = \begin{cases} s(1-t), & s \leq t, \\ t(1-s), & s > t. \end{cases} \quad (5.2)$$

In this context, we choose  $\mathcal{X} = \mathcal{Y} := L^2[0,1]$  such that the operator  $A$  is compact, selfadjoint and injective. Then the operator equation  $Ax = y$  can be rewritten as  $x = -y''$  provided that  $y \in H^2[0,1] \cap H_0^1[0,1]$ . Moreover, the operator  $A$  has the singular system (eigensystem)  $\{\sigma_j; u_j; u_j\}_{j=1}^\infty$  with  $\sigma_j = (j\pi)^{-1}$  and  $u_j(t) = \sqrt{2} \sin(j\pi t)$ . Furthermore, using the interpolation theory (see, e.g., [16]) it is not difficult to show that for  $4p - 1/2 \notin \mathbb{N}$

$$\mathcal{R}((A^*A)^p) = \left\{ x \in H^{4p}[0,1] : x^{2l}(0) = x^{2l}(1) = 0, l = 0, 1, \dots, [2p - 1/4] \right\}.$$

In our simulations, problem (5.2) is numerically solved by the linear finite element method. Let  $\mathcal{Y}_n$  be the finite element space of piecewise linear functions on a uniform grid with step size  $1/(n-1)$ . Denote by  $P_n$  the orthogonal projection operator acting from  $\mathcal{Y}$  into  $\mathcal{Y}_n$ . Define  $A_n := P_n A$  and  $\mathcal{X}_n := A_n^* \mathcal{Y}_n$ . Let  $\{\phi_j\}_{j=1}^n$  be a basis of  $\mathcal{Y}_n$ , then, instead of (5.2), we solve a system of linear equations  $A_n x_n = y_n$  in practice, where  $[A_n]_{ij} = \int_0^1 \left( \int_0^1 k(s,t)\phi_i(s)ds \right) \phi_j(t)dt$  and  $[y_n]_j = \int_0^1 y(t)\phi_j(t)dt$ .

We consider the following two different right-hand sides for (5.2).

EXAMPLE 5.1. In this example let  $y(s) := s(1-s)$ . Then the solution is  $x^\dagger \equiv 2$  and we have  $x^\dagger \in \mathcal{R}((A^*A)^p)$  for all  $p < 1/8$ .

EXAMPLE 5.2. In this example let  $y(s) := s^4(1-s)^3$ . Then the solution attains the form  $x^\dagger = -6t^2(1-t)(2-8t+7t^2)$ , and we have  $x^\dagger \in \mathcal{R}((A^*A)^p)$  for all  $p < 5/8$ .

Both examples are tested by using the grid size  $n = 100$ . Other parameters are:  $\Delta t = 19.4850$ ,  $x_0 = \dot{x}_0 = 0$ ,  $\tau = 3.1$ . Uniformly distributed noises with the magnitude  $\delta'$  are added to the discretized exact right-hand side:

$$[y_n^\delta]_j := [1 + \delta' \cdot (2\text{Rand}(x) - 1)] \cdot [y_n]_j, \quad j = 1, \dots, n,$$

where  $\text{Rand}(x)$  returns a pseudo-random value drawn from a uniform distribution on  $[0,1]$ . The noise level of measurement data is calculated by  $\delta = \|y_n^\delta - y_n\|_2$ , where  $\|\cdot\|_2$  denotes the standard vector norm in  $\mathbb{R}^n$ . The iteration of (5.1) is terminated according to the discrepancy principle ( $\tau = 3$  for all examples), i.e.

$$\|y^\delta - Ax_n^{k^*}\|_{L^2[0,1]} \leq \tau\delta < \|y^\delta - Ax_n^k\|_{L^2[0,1]}, \quad 0 \leq k < k^*.$$

Finally, to assess the accuracy of the approximate solutions, we define the  $L^2$ -norm relative error for an approximate solution  $x_n^{k^*}$ :  $\text{L2Err} := \|x_n^{k^*} - x^\dagger\|_{L^2[0,1]} / \|x^\dagger\|_{L^2[0,1]}$ , where  $x^\dagger$  is the exact solution to the corresponding model problem.

The results of the simulation are presented in Table 1, where we can conclude that, in general, the FAR method with  $\theta \in (1, 2)$  needs fewer iterations and offers more accurate regularized solutions. Concerning the number of iterations, the conjugate gradient method for the normal equation (CGNE, cf., e.g. [11]) performed significantly more effectively than all of other methods. However, the accuracy of the CGNE method is considerably worse than other accelerated regularization methods, since the step size of CGNE is too large to capture the optimal point and the semi-convergence effect disturbs the iteration rather early. Note that we set a maximal iteration number  $k_{max} = 200,000$  in all of our simulations.

## 6. Conclusion and outlook

In this paper, we have investigated the Fractional Asymptotical Regularization method (FAR) for solving linear ill-posed operator equations  $Ax = y$  with compact forward operators  $A$  mapping between infinite dimensional Hilbert spaces. Instead of exact right-hand side  $y$ , we are given noisy data  $y^\delta$  obeying the deterministic noise model  $\|y^\delta - y\| \leq \delta$ . We

TABLE 1. Comparisons with the Landweber method, Nesterov's method, Chebyshev method, and CGNE method.

$\delta$	$\theta = 0.5$		$\theta = 0.8$		Landweber	
	$k^*(\delta)$	L2Err	$k^*(\delta)$	L2Err	$k^*(\delta)$	L2Err
<b>Example 1:</b> $y(s) = s(1 - s)$						
0.01	4712	0.2694	1961	0.2684	712	0.2639
0.001	9336	0.1494	7057	0.1484	5997	0.1439
0.0001	$k_{max}$	0.2597	$k_{max}$	0.2073	$k_{max}$	0.1807
<b>Example 2:</b> $y(s) = s^4(1 - s)^3$						
0.001	3916	0.3038	2585	0.2918	1347	0.2833
0.0001	8534	0.2003	4947	0.1995	2034	0.1878
0.00001	$k_{max}$	0.3195	$k_{max}$	0.2184	$k_{max}$	0.1509
$\delta$	$\theta = 1.2$		$\theta = 1.5$		$\theta = 1.8$	
	$k^*(\delta)$	L2Err	$k^*(\delta)$	L2Err	$k^*(\delta)$	L2Err
<b>Example 1:</b> $y(s) = s(1 - s)$						
0.01	136	0.2476	29	0.2187	11	0.2204
0.001	733	0.1407	97	0.1321	45	0.1325
0.0001	8716	0.0997	1286	0.0673	657	0.0764
<b>Example 2:</b> $y(s) = s^4(1 - s)^3$						
0.001	224	0.2520	90	0.2019	58	0.2382
0.0001	883	0.0901	217	0.0767	155	0.0704
0.00001	4582	0.0287	328	0.0067	243	0.0083
$\delta$	Nesterov		Chebyshev		CGNE	
	$k^*(\delta)$	L2Err	$k^*(\delta)$	L2Err	$k^*(\delta)$	L2Err
<b>Example 1:</b> $y(s) = s(1 - s)$						
0.01	109	0.2590	127	0.2553	6	0.2213
0.001	704	0.1600	556	0.1496	19	0.1783
0.0001	4160	0.1025	4361	0.0897	38	0.0894
<b>Example 2:</b> $y(s) = s^4(1 - s)^3$						
0.001	102	0.2501	75	0.2434	6	0.2919
0.0001	208	0.1176	207	0.0990	13	0.1814
0.00001	1805	0.0280	2226	0.0196	15	0.0847

have proven that under both Hölder-type and logarithmic source conditions, FAR with fractional order  $\theta \in (0, 2)$  exhibits an optimal regularization method. Moreover, if  $\theta \in (1, 2)$ , FAR yields an accelerated method, namely, the optimal convergence rates can be obtained with approximately

the  $\theta$  root of iterations than needed for ordinary Landweber iteration. Finally, with the help of the one-step Adams-Moulton method, a new iterative regularization algorithm has been introduced.

As has been shown by this manuscript, fractional calculus can play some role for regularization schemes aimed at the stable approximate solution to general linear ill-posed problems. However, to the best of our knowledge, the literature in this direction is quite limited. The initial results in this work might be a bridge between fractional calculus and regularization theory. Of course, there are many remaining interesting problems in this topic. For instance, since the Hilbert scale can be understood as preconditioning for iterative regularization methods [4], it seems to be necessary to derive further assertions on acceleration of the FAR method in Hilbert scales. Certainly, the optimal choice of fractional order  $\theta$  and extensions to nonlinear problems are of interest. Yet, in the linear setting, new questions also arise, for example the study of the damped system

$$\left({}^C D_{0+}^\theta x^\delta\right)(t) + \eta \dot{x}^\delta(t) + A^* A x^\delta(t) = A^* y^\delta, \quad (6.3)$$

where  $\theta > 1$ , and  $\eta > 0$  is the damping parameter, which may or may not depend on the artificial time  $t$ . A remaining natural question is whether the damping term in (6.3) can yield a further acceleration of FAR.

### Acknowledgements

We express our gratitude to the anonymous referees whose valuable comments and suggestions allowed us to eliminate weak points of the manuscript and thus to improve the paper.

The work of Y. Zhang is supported by the Alexander von Humboldt Foundation through a postdoctoral researcher fellowship, and the work of B. Hofmann is supported by the German Research Foundation (DFG-grant HO 1454/12-1).

### References

- [1] V. Albani, P. Elbau, M. V. De Hoop and O. Scherzer. Optimal convergence rates results for linear inverse problems in Hilbert spaces. *Numer. Funct. Anal. Optim.* **37**, No 5 (2016), 521–540.
- [2] K. Diethelm and A. D. Freed. The FracPECE subroutine for the numerical solution of differential equations of fractional order. In *S. Heinzel and T. Plessner* (eds.): *Forschung und wissenschaftliches Rechnen: Beiträge zum Heinz-Billing Preis 1998*, Ges. f. wiss. Datenverarbeitung, Göttingen (1999), 57–71.
- [3] K. Diethelm, N.J. Ford and A.D. Freed. Detailed error analysis for a fractional Adams method. *Numer. Algorithms.* **36**, No 1 (2004), 31–52.

- [4] H. Egger and A. Neubauer. Preconditioning Landweber iteration in Hilbert scales. *Numer. Math.* **101**, No 4 (2005), 643–662.
- [5] H. Engl, M. Hanke and A. Neubauer. *Regularization of inverse problems*. Kluwer, Dordrecht (1996).
- [6] J. Flemming, B. Hofmann and P. Mathé. Sharp converse results for the regularization error using distance functions. *Inverse Problems* **27**, No 2 (2011), 025006 (18pp).
- [7] R. Gong, B. Hofmann and Y. Zhang. A new class of accelerated regularization methods, with application to bioluminescence tomography. *arXiv:1903.05972* (2019).
- [8] R. Gorenflo, A.A. Kilbas, F. Mainardi and S.V. Rogosin. *Mittag-Leffler Functions, Related Topics and Applications*. Springer-Verlag, Heidelberg (2014).
- [9] C.W. Groetsch. Integral equations of the first kind, inverse problems and regularization: a crash course. *J. Phys.: Conf. Ser.* **73** (2007), 012001.
- [10] P.C. Hansen. *Discrete Inverse Problems: Insight and Algorithms*. SIAM, Philadelphia (2010).
- [11] M. Hanke. *Conjugate gradient type methods for ill-posed problems*. Wiley, New York (1995).
- [12] B. Hofmann and P. Mathé. Analysis of profile functions for general linear regularization methods. *SIAM J. Numer. Anal.* **45**, No 3 (2007), 1122–1141.
- [13] T. Hohage. Regularization of exponentially ill-posed problems. *Numer. Funct. Anal. and Optimiz.* **21**, No 3-4 (2000), 439–464.
- [14] Y. Kian and M. Yamamoto. On existence and uniqueness of solutions for semilinear fractional wave equations. *Fract. Calc. Appl. Anal.* **20**, No 1 (2017), 117–138.
- [15] A.A. Kilbas, H.M. Srivastava and J.J. Trujillo. *Theory and Applications of Fractional Differential Equations. North-Holland Mathematics Studies, vol. 204*. Elsevier, Amsterdam (2006).
- [16] J. Lions and E. Magenes. *Non-Homogeneous Boundary Value Problems and Applications, Volumes I* Springer, Berlin (1972).
- [17] Z. Li, Y. Luchko and M. Yamamoto. Asymptotic estimates of solutions to initial-boundary value problems for distributed order time-fractional diffusion equations. *Fract. Calc. Appl. Anal.* **17**, No 4 (2014), 1114–1136.
- [18] Y. Liu, W. Rundell and M. Yamamoto. Strong maximum principle for fractional diffusion equations and an application to an inverse source problem. *Fract. Calc. Appl. Anal.* **19**, No 4 (2016), 888–906.

- [19] Y. Luchko and M. Yamamoto. General time-fractional diffusion equation: some uniqueness and existence results for the initial-boundary-value problems. *Fract. Calc. Appl. Anal.* **19**, No 3 (2016), 676–695.
- [20] F. Mainardi and R. Gorenflo. On Mittag-Leffler-type functions in fractional evolution processes. *J. Comput. Appl. Math.* **118** (2000), 283–299.
- [21] P. Mathé and S.V. Pereverzev. Geometry of linear ill-posed problems in variable Hilbert scales. *Inverse Problems* **19**, No 3 (2003), 789–803.
- [22] A. Neubauer. On Nesterov acceleration for Landweber iteration of linear ill-posed problems. *J. Inverse Ill-Posed Probl.* **25**, No 3 (2017), 381–390.
- [23] I. Podlubny. *Fractional Differential Equations*. Academic Press, San Diego (1999).
- [24] A. Rieder. Runge-Kutta integrators yield optimal regularization schemes. *Inverse Problems*. **21** (2005), 453–471.
- [25] T. Simon. Comparing Fréchet and positive stable laws. *Elec. J. Probab.* **19**, No 16 (2014), 1–25.
- [26] U. Tautenhahn. On the asymptotical regularization of nonlinear ill-posed problems. *Inverse Problems*. **10** (1994), 1405–1418.
- [27] G. Vainikko and A. Veretennikov. *Iteration Procedures in Ill-Posed Problems (In Russian)*. Nauka, Moscow (1986).
- [28] Y. Zhang and B. Hofmann. On the second-order asymptotical regularization of linear ill-posed inverse problems. *Appl. Anal.*, DOI:10.1080/00036811.2018.1517412 (2018).

<sup>1,2</sup> *Faculty of Mathematics, Chemnitz University of Technology  
Reichenhainer Str. 39/41, 09107 Chemnitz, Germany*

<sup>1</sup> *School of Science and Technology, Örebro University  
Fakultetsgatan 1, Örebro – 70182, Sweden*

<sup>1</sup> *e-mail: ye.zhang@mathematik.tu-chemnitz.de*

<sup>2</sup> *e-mail: hofmannb@mathematik.tu-chemnitz.de*