Conditional stability versus ill-posedness for operator equations with monotone operators in Hilbert space

Radu Ioan Boţ* and Bernd Hofmann†

September 16, 2016

Abstract

In the literature on singular perturbation (Lavrentiev regularization) for the stable approximate solution of operator equations with monotone operators in the Hilbert space the phenomena of conditional stability and local well-posedness or ill-posedness are rarely investigated. Our goal is to present some studies which try to bridge this gap. So we present new results on the impact of conditional stability on error estimates and convergence rates for the Lavrentiev regularization and distinguish for linear problems well-posedness and ill-posedness in a specific manner motivated by a saturation result. Taking into account that the behaviour of the bias (regularization error in the noise-free case) is crucial, general convergence rates, including logarithmic rates, are derived for linear operator equations by means of the method of approximate source conditions. This allows us to extend well-known convergence rates results for the Lavrentiev regularization that were based on general source conditions to the case of non-selfadjoint linear monotone forward operators for which general source conditions fail. Examples presenting the self-adjoint multiplication operator as well as the non-selfadjoint fractional integral operator and Cesàro operator illustrate the theoretical results. Extensions to the nonlinear case under specific conditions on the nonlinearity structure complete the paper.

MSC2010 subject classification: 47A52, 65F22, 47H05, 65J22, 65J15

Keywords: Conditional stability, local well-posedness, local ill-posedness, Hilbert space, equations with monotone operators, linear and nonlinear operator equations, Lavrentiev regularization, Tikhonov regularization, convergence rates, Hölder and logarithmic rates, approximate source conditions, solution smoothness, multiplication operator, fractional integration operator, Cesàro operator.

*Faculty of Mathematics, University of Vienna, Oskar-Morgenstern-Platz 1, A-1090 Vienna, Austria, Email: radu.bot@univie.ac.at .

†Faculty of Mathematics, Technische Universität Chemnitz, D-09107 Chemnitz, Germany, Email: bernd.hofmann@mathematik.tu-chemnitz.de .

1 Introduction

If $F : \mathcal{D}(F) \subseteq X \to Y$ denotes a sufficiently smooth and possibly nonlinear operator mapping between Hilbert spaces X and Y with norms $\|\cdot\|$, then it is not always trivial to find in a stable manner the solution $x^\dagger \in \mathcal{D}(F)$ to the operator equation

$$F(x) = y \tag{1.1}$$

with the exact right-hand side $y = F(x^\dagger)$ when only noisy data y^δ obeying the deterministic noise model

$$\|y - y^\delta\| \leq \delta \tag{1.2}$$

with noise level $\delta > 0$ are available. Even if (1.1) has $x^\dagger$ as the unique solution, a least squares approach

$$\|F(x) - y^\delta\|^2 \to \min, \qquad \text{subject to} \quad x \in \mathcal{D}(F),$$

is not always successful if the Hilbert space is infinite dimensional. Then the least squares minimizers need not exist and if they exist their convergence to $x^\dagger$ in the norm of X as $\delta \to 0$ can only be expected if the operator equation is locally well-posed at $x^\dagger$. In this context, we recall the following definition introduced in [25, Definition 2].

Definition 1.1. *The equation (1.1) is called locally well-posed at the solution point $x^\dagger \in \mathcal{D}(F)$ if there is a ball $\mathcal{B}_r(x^\dagger)$ with radius $r > 0$ and center $x^\dagger$ such that for every sequence $\{x_k\}_{k=1}^\infty \subset \mathcal{B}_r(x^\dagger) \cap \mathcal{D}(F)$ the convergence of images $\lim_{k\to\infty} \|F(x_k) - F(x^\dagger)\| = 0$ implies the convergence of the preimages $\lim_{k\to\infty} \|x_k - x^\dagger\| = 0$. Otherwise it is called locally ill-posed.*

In particular if the equation (1.1) is a model of an inverse and therefore mostly ill-posed problem, it makes sense to exploit a singularly perturbed auxiliary problem to equation (1.1), which is automatically locally well-posed. The most prominent such approach is the Tikhonov regularization, where in the simplest case (cf. [13, Chapt. 10]) stable approximate solutions $x_\alpha^\delta \in \mathcal{D}(F)$ solve the extremal problem

$$\|F(x) - y^\delta\|^2 + \|x - \bar{x}\|^2 \to \min, \qquad \text{subject to} \quad x \in \mathcal{D}(F), \tag{1.3}$$

with regularization parameter $\alpha > 0$ and reference element (initial guess) $\bar{x} \in X$. Variants of Tikhonov regularization, however, are also helpful and advantageous (cf., e.g., [10, 11] and [26, §6.2]) if (1.1) is locally well-posed in the sense that a *conditional stability estimate* of the form

$$\|x - x^\dagger\| \leq \varphi(\|F(x) - F(x^\dagger)\|) \qquad \text{for all} \quad x \in \mathcal{D}(F) \cap Q \tag{1.4}$$

applies, with some set $Q \subset X$ containing $x^\dagger$ and some *concave* index function φ, where we call $\varphi : [0, \infty) \to [0, \infty)$ *index function* if it is continuous, strictly increasing and satisfies the condition $\varphi(0) = 0$. Then the method ensures convergence and rates of the approximate solutions when the regularized solutions are embedded in the stability region $\mathcal{D}(F) \cap Q$.

The focus of this paper is on the specific situation of an operator equation (1.1) with $Y = X$, $\mathcal{D}(F) = X$, and *monotone* operators F as characterized by the following assumption. Mostly, *infinite dimensional* Hilbert spaces X will be under consideration, but for examples also finite dimensional cases shall be exploited.

Assumption 1.2. *Consider the operator equation (1.1) with solution $x^\dagger \in X$ under the auspices that*

(a) X is a real separable Hilbert space with norm $\|\cdot\|$ and inner product $\langle\cdot,\cdot\rangle$ and

(b) $F : X \to X$ is a monotone operator, i.e.

$$\langle F(x) - F(\tilde{x}), x - \tilde{x}\rangle \geq 0 \qquad \text{for all} \quad x, \tilde{x} \in X, \tag{1.5}$$

which is moreover hemicontinuous and hence maximally monotone.

Under Assumption 1.2 there occur well-posed and ill-posed situations. The best situation of global well-posedness is characterized by *strong monotonicity*

$$\langle F(x) - F(\tilde{x}), x - \tilde{x}\rangle \geq C \|x - \tilde{x}\|^2 \qquad \text{for all} \quad x, \tilde{x} \in X, \tag{1.6}$$

with some constant $C > 0$, which implies the *coercivity* condition

$$\lim_{\|x\|\to\infty} \frac{\langle F(x), x\rangle}{\|x\|} = \infty. \tag{1.7}$$

Proposition 1.3. *Under the requirements of Assumption 1.2 strengthened by the condition (1.6) the equation (1.1) is uniquely solvable in X for all $y \in X$, and the solutions are Lipschitz continuous with respect to the data, i.e. the inverse operator $F^{-1} : X \to X$ is well-defined with*

$$\|F^{-1}(y) - F^{-1}(\tilde{y})\| \leq \frac{1}{C} \|y - \tilde{y}\| \qquad \text{for all} \quad y, \tilde{y} \in X. \tag{1.8}$$

Proof. The Browder-Minty theorem ensures under the supposed conditions that F is surjective and due to (1.6) even bijective. Also from (1.6) we have for all $x, \tilde{x} \in X$

$$C \|x - \tilde{x}\|^2 \leq \langle F(x) - F(\tilde{x}), x - \tilde{x}\rangle \leq \|x - \tilde{x}\| \, \|F(x) - F(\tilde{x})\|,$$

which yields (1.8) and completes the proof. □

However, there are a lot of examples for inverse problems occurring in natural sciences, engineering, and finance, where estimates of the form (1.8) fail and operator equations (1.1) with monotone forward operators F have to be solved in a stable approximate manner. Due to the smoothing character of F in these cases, local ill-posedness must be conjectured, and we have so-called operator equations of the first kind. Compact monotone operators F are typical for that situation. Then F obeys (1.5), but fails to satisfy inequalities of the form (1.6). Examples of ill-posed problems in integral and differential equations under monotone forward operators are, for example, presented in [1, Section 1.3] and [23, Section 5]. However, for all $\alpha > 0$ the associated equations of the second kind $G(x) = y$ with $G(x) := F(x) + \alpha\, x$ possess a strongly monotone forward operator G and are hence locally well-posed everywhere, because we have

$$\langle G(x) - G(\tilde{x}), x - \tilde{x}\rangle = \langle F(x) - F(\tilde{x}), x - \tilde{x}\rangle + \alpha \|x - \tilde{x}\|^2 \geq \alpha \|x - \tilde{x}\|^2 \ \text{ for all } \ x, \tilde{x} \in X.$$

This gives a substantial motivation for using singular perturbations for the stable approximate solution of equation (1.1) also here. Due to the maximal monotonicity of F the simpler Lavrentiev regularization (cf. the seminal monograph [28] as well as the more recent works [1, 29]) is applicable, where stable approximate solutions $x_\alpha^\delta \in X$ solve the operator equation

$$F(x_\alpha^\delta) + \alpha(x_\alpha^\delta - \bar{x}) = y^\delta, \tag{1.9}$$

with regularization parameter $\alpha > 0$ and reference element $\bar{x} \in X$. Such approach is also helpful if Proposition 1.3 is not applicable, because coercivity (1.7) fails or well-posedness at $x^\dagger$ takes place only in a local sense. The latter is the case if F is *strongly monotone* in a neighbourhood of $x^\dagger$, i.e. the locally relaxed version of (1.6),

$$\langle F(x) - F(x^\dagger), x - x^\dagger \rangle \geq C \, \|x - x^\dagger\|^2 \qquad \text{for all} \quad x \in \mathcal{B}_r(x^\dagger), \tag{1.10}$$

with some radius $r > 0$ and some constant $C > 0$ is valid, or if F is *uniformly monotone* in a neighbourhood of $x^\dagger$, i.e.

$$\langle F(x) - F(x^\dagger), x - x^\dagger \rangle \geq \zeta(\|x - x^\dagger\|) \qquad \text{for all} \quad x \in \mathcal{B}_r(x^\dagger) \tag{1.11}$$

holds with some radius $r > 0$ and some index function ζ. In both situations we have local well-posedness at $x^\dagger$, and we refer for examples to the monograph [42] and also to papers like [6].

Theorem 1.4. *Let the inequality (1.11) hold with an index function ζ of the form $\zeta(t) = \theta(t)\, t$, $t > 0$, such that θ is a convex index function. Then the condition (1.11) of local uniform monotonicity is a conditional stability estimate of the form (1.4) with $\mathcal{D}(F) = X$, $Q = \mathcal{B}_r(x^\dagger)$, and the concave index function $\varphi(t) = \theta^{-1}(t)$, $t > 0$. This implies that the operator equation (1.1) is locally well-posed at the solution point $x^\dagger$. Evidently, these assertions apply for the local strong monotonicity (1.10) yielding (1.4) with $\varphi(t) = \frac{1}{C}\, t$.*

Proof. For $x \in \mathcal{B}_r(x^\dagger)$ we can estimate from (1.11) as

$$\zeta(\|x - x^\dagger\|) = \theta(\|x - x^\dagger\|)\, \|x - x^\dagger\| \leq \langle F(x) - F(x^\dagger), x - x^\dagger \rangle \leq \|F(x) - F(x^\dagger)\|\, \|x - x^\dagger\|$$

and hence $\|x - x^\dagger\| \leq \theta^{-1}(\|F(x) - F(x^\dagger)\|)$, where θ^{-1} is a concave index function which plays the role of φ in (1.4). The special case of local strong monotonicity (1.10) applies here with $\theta(t) = C\, t$. □

Remark 1.5. In very specific cases, see Example 2.4 below, the function θ in Theorem 1.4 can also be concave such that θ^{-1} is a convex index function. Then, surprisingly, with $0 < \kappa < 1$ the convergence rate (2.8) in Corollary 2.3 below can be overlinear as $\mathcal{O}(\delta^{1/\kappa})$.

A special case of (1.1) taking into account Assumption 1.2 is characterized by forward operators $A \in \mathcal{L}(X)$ instead of F, where $\mathcal{L}(X)$ denotes the Banach space of bounded linear operators $A : X \to X$ and $\|A\|$ indicates the corresponding operator norm. So we consider in this case linear operator equations

$$Ax = y \tag{1.12}$$

under the noise model (1.2), where A is monotone (accretive), i.e.

$$\langle Ax, x \rangle \geq 0 \qquad \text{for all} \quad x \in X. \tag{1.13}$$

Note that for all such operators A and all $\alpha > 0$ the properties

$$(A + \alpha I)^{-1} \in \mathcal{L}(X), \qquad \|(A + \alpha I)^{-1} A\| \leq 1, \tag{1.14}$$

and

$$\|(A + \alpha I)^{-1}\| \leq \frac{1}{\alpha} \tag{1.15}$$

are valid (cf. [18, Section 7.1.1]. Regularized solutions x_α^δ of Lavrentiev regularization in the linear case attain the explicit form

$$x_\alpha^\delta = (A + \alpha I)^{-1}(y^\delta + \alpha\, \bar{x}), \tag{1.16}$$

because they solve the equation

$$A x_\alpha^\delta + \alpha(x_\alpha^\delta - \bar{x}) = y^\delta. \tag{1.17}$$

Since the properties of a linear operator $A \in \mathcal{L}(X)$ do not depend on the solution point $x^\dagger$, well-posedness and ill-posedness of the operator equation (1.12) in the sense of Definition 1.1 are global properties. Thus, the equation is locally well-posed everywhere or locally ill-posed everywhere as the following proposition outlines.

Proposition 1.6. *The linear operator equation (1.12) is under (1.13) locally well-posed everywhere if and only if A is continuously invertible, i.e. if $A^{-1} \in \mathcal{L}(X)$ and we have a constant $K > 0$ such that*

$$\|(A + \alpha I)^{-1}\| \leq K < \infty \qquad \textit{for all} \quad \alpha > 0, \tag{1.18}$$

where $K = \|A^{-1}\|$ holds true. Alternatively, (1.13) is locally ill-posed everywhere if and only if the nullspace of A is non-trivial, i.e. $\mathcal{N}(A) \neq \{0\}$, or the range $R(A)$ of A is not closed. Then we have

$$\|(A + \alpha I)^{-1}\| = \frac{1}{\alpha} \qquad \textit{for all} \quad \alpha > 0. \tag{1.19}$$

Proof. The well-posed case (1.18) is characterized by $0 \notin \sigma(A)$, where $\sigma(A)$ denotes the spectrum of the operator A, whereas the ill-posed case is characterized by $0 \in \sigma(A)$.

For $0 \notin \sigma(A)$ we have by definition $A^{-1} \in \mathcal{L}(X)$, i.e. with bijective operator $A : X \to X$, $\|A^{-1}\| < \infty$ and $\|A^{-1}(y - \tilde{y})\| \leq \|A^{-1}\|\, \|y - \tilde{y}\|$ for all $y, \tilde{y} \in X$, which indicates the local well-posedness everywhere. With A also A^{-1} is monotone and thus we can estimate with $K := \|A^{-1}\|$ as

$$\|(A + \alpha I)^{-1}\| = \|A^{-1}(I + \alpha A^{-1})^{-1}\| \leq \|A^{-1}\|\, \|(I + \alpha A^{-1})^{-1}\| \leq K \quad \text{for all} \quad \alpha > 0.$$

For $0 \in \sigma(A)$ we have by definition that at least one of the assertions $N(A) \neq \{0\}$ and $R(A)$ is not closed is true, which can be summarized by the condition $\mathcal{R}(A) \neq X$ due to the orthogonal sum

$$\overline{\mathcal{R}(A)} \oplus \mathcal{N}(A) = X \tag{1.20}$$

(cf., e.g., [34, Theorem 1.1.10]). Both cases indicate local ill-posedness everywhere. This is obvious for $\mathcal{N}(A) \neq \{0\}$. In the case $\mathcal{N}(A) = \{0\}$, but $\mathcal{R}(A) \neq \overline{\mathcal{R}(A)}$, we have that A^{-1}

exists and is an unbounded linear operator and hence for all $r > 0$ that there is sequence $\{x_n\} \subset X$ with $\|x_n\| = r$ and $\lim_{n\to\infty} \|Ax_n\| = 0$. Then, we have for $x^\dagger + x_n \in \mathcal{B}_r(x^\dagger)$ the limit properties $x^\dagger + x_n \not\to x^\dagger$ but $A(x^\dagger + x_n) \to Ax^\dagger$ as $n \to \infty$ and thus local ill-posedness at $x^\dagger$. Taking into account (1.15), to prove (1.19) it remains to show $\|(A+\alpha I)^{-1}\| \geq \frac{1}{\alpha}$ for all $\alpha > 0$. This, however, is a consequence of the Neumann series theory, which says that we have, for a bounded linear operator $B : X \to X$ with $\|B\| < 1$, that $(I - B)^{-1} \in \mathcal{L}(X)$. By setting $B := (\frac{1}{\alpha}A + I)^{-1}$, we must have $\|B\| \geq 1$. Otherwise, we would get that $0 \notin \sigma(A)$. Evidently, the conditions (1.18) and (1.19) are incompatible, but one of them is always true for a bounded monotone operator A. Now the proof is complete. □

Note that the ill-posed case in Proposition 1.6 with $\mathcal{R}(A) \neq \overline{\mathcal{R}(A)}$ can only occur if X is an infinite dimensional space and the range $\mathcal{R}(A)$ is also infinite dimensional. Moreover, it will be a by-product of the assertion of Proposition 3.4 (cf. (3.14)) below that for arbitrary monotone operators $A \in \mathcal{L}(X)$ the condition

$$\|(A + \alpha I)^{-1}x\| = O(1) \qquad \text{as} \qquad \alpha \to 0,$$

which is in the case (1.18) valid for all $x \in X$, cannot be improved to

$$\|(A + \alpha I)^{-1}x\| = o(1) \qquad \text{as} \qquad \alpha \to 0$$

if $x \neq 0$.

It is evident that strong monotonicity

$$\langle Ax, x\rangle \geq C\,\|x\|^2 \qquad \text{for all} \quad x \in X, \tag{1.21}$$

with some constant $C > 0$ implies $\|x\| \leq \frac{1}{C}\|Ax\|$ for all $x \in X$ and hence with $A^{-1} \in \mathcal{L}(X)$ local well-posedness of (1.13) everywhere. Vice versa, $A^{-1} \in \mathcal{L}(X)$ does not, in general, imply strong monotonicity, because we have $\langle Ax, x\rangle = 0$ for all $x \in X$ if the monotone operator A is skew-symmetric, i.e. for the adjoint operator A^* that $A^* = -A$. The simplest case of such behaviour is $A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ for $X = \mathbb{R}^2$.

Remark 1.7. For linear equations (1.12) with monotone $A \in \mathcal{L}(X)$, the case distinction (cf. Proposition 1.6) between locally well-posed and ill-posed situations based on Definition 1.1 is different from the usual case distinction in the literature of linear regularization theory (see, e.g., [33]), where a bounded pseudoinverse $A^\dagger$ characterized by $\mathcal{R}(A) = \overline{\mathcal{R}(A)}$ denotes well-posedness and an unbounded $A^\dagger$ characterized by $\mathcal{R}(A) \neq \overline{\mathcal{R}(A)}$ denotes ill-posedness. However, we will see below in Proposition 2.6 that the concept of Definition 1.1 is the more appropriate one for our setting in the context of Lavrentiev regularization.

The remaining part of the paper is organized as follows: In Section 2, we discuss the impact of conditional stability on error estimates and convergence rates for the Lavrentiev regularization. Furthermore, we mention in Proposition 2.6 some saturation result from [35] for the linear case which motivates to distinguish well-posedness and ill-posedness on the basis of Definition 1.1. The role of the regularization error in the noise-free case, called bias, will be investigated in Section 3 for nonlinear and linear problems. For linear operator equations general convergence rates, including logarithmic rates, are derived in

Section 4 by means of the method of approximate source conditions. This allows us to extend well-known convergence rates results for the Lavrentiev regularization, which were based on general source conditions, to the case of non-selfadjoint linear monotone forward operators for which general source conditions fail. Examples presenting the self-adjoint multiplication operator as well as the non-selfadjoint fractional integral operator and Cesàro operator illustrate the theoretical results of this section. Extensions to the nonlinear case under specific conditions on the nonlinearity structure in Section 5 complete the paper.

2 Error estimates and the case of conditional stability

Proposition 2.1. *Let under the Assumption 1.2 the solution set to equation (1.1)*

$$L := \{x \in X : \ F(x) = y\}$$

be nonempty. This set L is closed and convex, and consequently there is a uniquely determined $\bar{x}$-minimum norm solution $x^\dagger_{mn} \in L$ to (1.1) such that

$$\|x^\dagger_{mn} - \bar{x}\| = \min\{\|x^\dagger - \bar{x}\| : \ x^\dagger \in L\}.$$

The Lavrentiev-regularized solution $x^\delta_\alpha \in X$ is uniquely determined, which means that (1.9) has a unique solution x^δ_α for all $\bar{x} \in X$, $y^\delta \in X$ and $\alpha > 0$, which depends continuously on y^δ. Moreover, for any solution $x^\dagger \in L$, the following three basic inequalities are valid:

$$\|x^\delta_\alpha - x^\dagger\|^2 \leq \langle x^\dagger - \bar{x}, x^\dagger - x^\delta_\alpha\rangle + \frac{\delta}{\alpha}\|x^\delta_\alpha - x^\dagger\|, \tag{2.1}$$

$$\|x^\delta_\alpha - x^\dagger\| \leq \|x^\dagger - \bar{x}\| + \frac{\delta}{\alpha}, \tag{2.2}$$

$$\|F(x^\delta_\alpha) - F(x^\dagger)\| \leq \alpha\|x^\dagger - \bar{x}\| + \delta. \tag{2.3}$$

Proof. The closedness and convexity of L is due to the maximal monotonicity of $F : X \to X$ (cf. [5, Prop. 23.39]). Then the $\bar{x}$-minimum norm solution $x^\dagger_{mn}$ is the uniquely determined best approximation of $\bar{x}$ in L. The next assertion of the proposition is a consequence of the Browder-Minty theorem which ensures that for all $\alpha > 0$ the operator $F + \alpha I : X \to X$ is bijective and strongly monotone such that x^δ_α is uniquely determined and depends continuously on the data y^δ. As outlined in [23], by testing (1.9) with the two elements $x^\delta_\alpha - x^\dagger$ and $F(x^\delta_\alpha) - F(x^\dagger)$ we obtain

$$\begin{aligned}&\langle F(x^\delta_\alpha) - F(x^\dagger), x^\delta_\alpha - x^\dagger\rangle + \langle y - y^\delta, x^\delta_\alpha - x^\dagger\rangle \\ &+\alpha\|x^\delta_\alpha - x^\dagger\|^2 + \alpha\langle x^\dagger - \bar{x}, x^\delta_\alpha - x^\dagger\rangle = 0\end{aligned} \tag{2.4}$$

and

$$\begin{aligned}&\|F(x^\delta_\alpha) - F(x^\dagger)\|^2 + \langle y - y^\delta, F(x^\delta_\alpha) - F(x^\dagger)\rangle \\ &+\alpha\langle F(x^\delta_\alpha) - F(x^\dagger), x^\delta_\alpha - x^\dagger\rangle + \alpha\langle x^\dagger - \bar{x}, F(x^\delta_\alpha) - F(x^\dagger)\rangle = 0\,,\end{aligned} \tag{2.5}$$

respectively. By using the monotonicity (1.5) of F, the Cauchy-Schwarz inequality yields (2.1) and moreover (2.2), as a consequence of (2.4), while (2.3) follows as a consequence of (2.5). This completes the proof. □

Theorem 2.2. *For a solution $x^\dagger$ of equation (1.1), assume that a conditional stability estimate of the form (1.4) with $\mathcal{D}(F) = X,\ Q = \mathcal{B}_r(x^\dagger)$, and some concave index function φ holds. Then the solution set L is a singleton, i.e. $L = \{x^\dagger\}$, and for an a priori parameter choice $\alpha = \alpha(\delta) = c\,\delta,\ c > 0$, the Lavrentiev regularized solutions x_α^δ convergence to $x^\dagger$ with the rate*

$$\|x_{\alpha(\delta)}^\delta - x^\dagger\| = \mathcal{O}(\varphi(\delta)) \qquad as \qquad \delta \to 0 \tag{2.6}$$

if the radius r in (1.4) is sufficiently large such that $r > \|x^\dagger - \bar{x}\| + \frac{1}{c}$.

Proof. A conditional stability estimate of the form (1.4) with $\mathcal{D}(F) = X,\ Q = \mathcal{B}_r(x^\dagger)$ and some $r > 0$ ensures that $L \cap \mathcal{B}_r(x^\dagger) = \{x^\dagger\}$ is a singleton. As L is a convex set (cf. Proposition 2.1), there cannot be a second element in L. Now we have from (2.2) that $\|x_{\alpha(\delta)}^\delta - x^\dagger\| \le \|x^\dagger - \bar{x}\| + \frac{1}{c}$ and thus $x_{\alpha(\delta)}^\delta \in \mathcal{B}_r(x^\dagger)$. Then (1.4) and (2.3) yield

$$\|x_{\alpha(\delta)}^\delta - x^\dagger\| \le \varphi(\|F(x_{\alpha(\delta)}^\delta) - F(x^\dagger)\|) \le \varphi(c\,\delta\,\|x^\dagger - \bar{x}\| + \delta) \le 2\,\max(1, c\,\|x^\dagger - \bar{x}\|)\,\varphi(\delta).$$

This proves the theorem. □

As was mentioned above for the Tikhonov regularization, Theorem 2.2 shows that also the Lavrentiev regularization ensures convergence and rates of the approximate solutions under the conditional stability estimate (1.4) by embedding the regularized solutions in the stability region, which is here $\mathcal{B}_r(x^\dagger)$. From Theorems 1.4 and 2.2 we immediately arrive at the following corollary.

Corollary 2.3. *Choose the regularization parameter for the Lavrentiev regularization a priori as $\alpha(\delta) \sim \delta$. If F is strongly monotone with sufficiently large $r > 0$ in (1.10), then we have a linear (Lipschitz) convergence rate*

$$\|x_{\alpha(\delta)}^\delta - x^\dagger\| = \mathcal{O}(\delta) \qquad as \qquad \delta \to 0. \tag{2.7}$$

If F is uniformly monotone with $\zeta(t) = t^{\kappa+1},\ \kappa > 1$, and sufficiently large $r > 0$ in (1.11), then we have a Hölder convergence rate

$$\|x_{\alpha(\delta)}^\delta - x^\dagger\| = \mathcal{O}(\delta^{1/\kappa}) \qquad as \qquad \delta \to 0. \tag{2.8}$$

Example 2.4. (One dimensional example) For $X := \mathbb{R}$ with $\|x\| := |x|$ we consider the continuous monotone operator $F : \mathbb{R} \to \mathbb{R}$ defined for exponents $\kappa > 0$ as

$$F(x) := \begin{cases} -1 & \text{if} \quad -\infty < x < -1 \\ -(-x)^\kappa & \text{if} \quad -1 \le x \le 0 \\ x^\kappa & \text{if} \quad 0 < x \le 1 \\ 1 & \text{if} \quad 1 < x < \infty \end{cases},$$

which however is not bijective and not coercive. Then we have obviously local ill-posedness at $x^\dagger$ if $x^\dagger < -1$ or $x^\dagger > 1$. On the other hand we have for all $\kappa > 0$ local well-posedness at $x^\dagger = 0$, because the local uniform monotonicity condition (1.11) is satisfied there with $\zeta(t) = t^{\kappa+1}$ such that Theorem 1.4 and Corollary 2.3 apply for $x^\dagger = 0$ with $\theta(t) = t^\kappa$ for $\kappa \ge 1$. Indeed, a superlinear convergence rate (2.8) at $x^\dagger = 0$ occurs if $0 < \kappa < 1$.

Remark 2.5. As fixed in item (b) of Assumption 1.2 we restrict our considerations throughout this paper in a conscious way to the case $\mathcal{D}(F) = X$, where the operator F is monotone on the whole Hilbert space X. By contrast, Tautenhahn [40] and subsequently other authors apply and analyze the Lavrentiev regularization (1.9) for Fréchet differentiable nonlinear operators F which are only defined and monotone on a ball $\mathcal{B}_r(x^\dagger)$ with sufficiently large radius

$$r > \|x^\dagger - \bar{x}\| + \delta/\alpha. \tag{2.9}$$

Then regularized solutions x_α^δ can only exist in a systematic way if for fixed $\delta > 0$ the regularization parameter $\alpha > 0$ is not too small. Otherwise, even if δ and α are fixed such that (2.9) is satisfied, to apply Lavrentiev regularization it must be shown that for arbitrary data y^δ with (1.2) solutions x_α^δ to equation (1.9) always exist. At the moment this is not evident for the authors, because we cannot retrace the arguments in the proof sketch of Theorem 1.1 in [40].

For the special case of monotone linear operators $A \in \mathcal{L}(X)$ we have the two different situations formulated in Proposition 2.6. This indicates a significant gap in the convergence rates and motivates the specific case distinction between well-posedness and ill-posedness based on Definition 1.1 also for linear monotone operators A as mentioned above in Remark 1.7.

Proposition 2.6. *For the maximal best possible error*

$$E_{x^\dagger}(\delta) := \sup_{y^\delta \in X:\, \|y - y^\delta\| \le \delta} \inf_{\alpha > 0} \|x_\alpha^\delta - x^\dagger\|$$

of Lavrentiev regularization to equation (1.12) with bounded monotone linear operator A we have on the one hand

$$E_{x^\dagger}(\delta) = \mathcal{O}(\delta) \qquad as \qquad \delta \to 0 \tag{2.10}$$

for all $x^\dagger \in X$ if (1.18) is valid, i.e. if A is continuously invertible. On the other hand, we have that

$$E_{x^\dagger}(\delta) = o(\sqrt{\delta}) \qquad as \qquad \delta \to 0 \qquad implies \qquad x^\dagger - \bar{x} = 0 \tag{2.11}$$

if (1.18) is violated for arbitrarily large $K > 0$, that is exactly the case if the null-space $\mathcal{N}(A)$ of A is not trivial or the range $\mathcal{R}(A)$ of A is not closed.

Proof. For the special case of monotone linear operators $A \in \mathcal{L}(X)$ we find directly from (1.16) the error estimate

$$\|x_\alpha^\delta - x^\dagger\| \le \|(A + \alpha I)^{-1}[(y^\delta - y) + \alpha(\bar{x} - x^\dagger)]\| \le \|(A + \alpha I)^{-1}\|(\delta + \alpha \|x^\dagger - \bar{x}\|). \tag{2.12}$$

This ensures the linear convergence rate (2.7) for the regularization parameter choice $\alpha(\delta) \sim \delta$ in the well-posed case (1.18) and hence (2.10). The implication (2.11), however, recalls the recently published saturation result from Theorem 5.1 in [35] for the ill-posed case. This proves a significant gap in the convergence rates between well-posed and ill-posed situations. □

It should be mentioned that the saturation result (2.11) for noisy data is a Lavrentiev regularization analogue to the well-known saturation result [13, Proposition 5.3] for the Tikhonov regularization.

3 The distinguished role of bias

For the error analysis of Lavrentiev-regularized solutions it is helpful to consider in addition to x_α^δ the regularized solutions $x_\alpha := x_\alpha^0$ in the noise-free case ($\delta = 0$), which satisfy the operator equations

$$F(x_\alpha) + \alpha(x_\alpha - \bar{x}) = y \tag{3.1}$$

and

$$Ax_\alpha + \alpha(x_\alpha - \bar{x}) = y \tag{3.2}$$

in the nonlinear and linear case, respectively. From Proposition 2.1 we have that also the elements $x_\alpha \in X$ are uniquely determined for all $\alpha > 0$. It is obvious in regularization theory that the total norm error of regularization can be estimated above by the triangle inequality as

$$\|x_\alpha^\delta - x^\dagger\| \leq \|x_\alpha - x^\dagger\| + \|x_\alpha - x_\alpha^\delta\| \tag{3.3}$$

such that an upper bound of the noise propagation error $\|x_\alpha - x_\alpha^\delta\|$ is independent of the solution $x^\dagger$, but depends on the noise level δ and on the regularization parameter $\alpha > 0$. If the regularization procedure R_α is expressed by a continuous linear mapping $y^\delta \mapsto x_\alpha^\delta$, then estimates of the form $\|x_\alpha - x_\alpha^\delta\| \leq \|R_\alpha\|\delta$ are standard, where $\lim_{\alpha\to 0} \|R_\alpha\| = \infty$ takes place for the ill-posed case. So we have for (3.2), in the case $\bar{x} = 0$, $R_\alpha y = (A + \alpha I)^{-1} y$ with

$$\|x_\alpha - x_\alpha^\delta\| \leq \frac{\delta}{\alpha} \tag{3.4}$$

for all $\delta \geq 0$ and $\alpha > 0$ due to (1.15). For nonlinear ill-posed problems, however, the regularization procedure is in general characterized by a nonlinear mapping $y^\delta \mapsto x_\alpha^\delta$. Estimates of $\|x_\alpha - x_\alpha^\delta\|$ from above independent of $x^\dagger$ are then restricted to classes of forward operator F with specific nonlinearity properties, and we refer for example to the discussion in [37] for estimates of the form $\|x_\alpha - x_\alpha^\delta\| \leq \frac{c\delta}{\sqrt{\alpha}}$ versus $\|x_\alpha - x_\alpha^\delta\| \leq \frac{c\delta}{\alpha}$ for the nonlinear Tikhonov regularization (cf. (1.3)). Taking advantage of the monotonicity of F the situation is simpler for the nonlinear Lavrentiev regularization as the following lemma shows.

Lemma 3.1. *For arbitrary monotone forward operators F and solutions $x^\dagger \in X$ we have the uniform propagation error estimate (3.4) for the Lavrentiev regularization with x_α^δ and x_α from (1.9) and (3.1), respectively.*

Proof. For $x_\alpha^\delta - x_\alpha = 0$, (3.4) is trivially satisfied. As difference of the two equations (1.9) and (3.1) we have the equation

$$F(x_\alpha^\delta) - F(x_\alpha) + \alpha(x_\alpha^\delta - x_\alpha) = y^\delta - y$$

and thus by testing with $x_\alpha^\delta - x_\alpha \neq 0$

$$\langle F(x_\alpha^\delta) - F(x_\alpha), x_\alpha^\delta - x_\alpha\rangle + \alpha \|x_\alpha^\delta - x_\alpha\|^2 = \langle y^\delta - y, x_\alpha^\delta - x_\alpha\rangle \leq \|x_\alpha^\delta - x_\alpha\|\,\delta.$$

Due to the monotonicity property (1.5) this implies the inequality $\alpha\|x_\alpha^\delta - x_\alpha\| \leq \delta$ and hence (3.4) which completes the proof. □

Taking into account the bound (3.4) of the noise propagation error, for fixed forward operator F and fixed solution $x^\dagger$ the asymptotics of the total regularization error $\|x_\alpha^\delta - x^\dagger\|$ as $\delta \to 0$ of Lavrentiev regularization is essentially influenced by the asymptotics of the bias (regularization error for noise-free data)

$$B_{x^\dagger}^F(\alpha) := \|x_\alpha - x^\dagger\|$$

as $\alpha \to 0$. From [5, Section 23] we derive the following proposition for the bias of the Lavrentiev regularization.

Proposition 3.2. *Under Assumption 1.2 let $x^\dagger$ solve the equation (1.1). Then we have*

$$\lim_{\alpha \to 0} B_{x^\dagger}^F(\alpha) = 0 \tag{3.5}$$

if and only if $x^\dagger = x_{mn}^\dagger \in X$, i.e. $x^\dagger$ is the $\bar{x}$-minimum norm solution to equation (1.1).

Proof. Under the stated assumptions we have from Proposition 2.1 that the nonempty set $L := \{x \in X : F(x) = y\}$ is closed and convex and hence the projection of $\bar{x}$ onto this set is an $\bar{x}$-minimum norm solution to equation (1.1) and uniquely determined. Then we have from Theorem 23.44 (i) in [5] that the uniquely determined element $x_\alpha \in X$ existing for all $\alpha > 0$ and $\bar{x} \in X$, which satisfies the equation

$$F(x_\alpha) - y + \alpha(x_\alpha - \bar{x}) = 0,$$

tends in the norm of X to the projection of $\bar{x}$ on L. If $x^\dagger$ solves the equation (1.1), but fails to be an $\bar{x}$-minimum norm solution, then the projection of $\bar{x}$ on L differs from $x^\dagger$ and (3.5) cannot hold. This proves the proposition. □

The asymptotic behaviour of the bias $B_{x^\dagger}^F(\alpha) \to 0$ as $\alpha \to 0$ expressing the intrinsic smoothness of the solution $x^\dagger$ with respect to the forward operator F (cf. in a more general context the ideas in [38, Chapt. 3] and [22]) fully determines the specific error profile for the solution $x^\dagger$. Therefore the bias was called 'profile function' in the former paper [24] with focus on a general regularization scheme for linear ill-posed problems.

Remark 3.3. For the Lavrentiev regularization (1.17) to linear problems (1.12) with monotone forward operator $A \in \mathcal{L}(X)$, Proposition 3.2 applies and based on formula (1.20) we note that a solution $x^\dagger$ to the equation (1.12) is an $\bar{x}$-minimum solution $x_{mn}^\dagger$ if and only if $x^\dagger - \bar{x}$ is orthogonal to the null-space of the linear operator A, i.e.

$$x^\dagger - \bar{x} \perp \mathcal{N}(A). \tag{3.6}$$

The corresponding bias attains the form

$$B_{x^\dagger}^A(\alpha) = \|x_\alpha - x^\dagger\| = \alpha \, \|(A + \alpha I)^{-1}(x^\dagger - \bar{x})\| \tag{3.7}$$

and we have $\lim_{\alpha \to 0} B_{x_{mn}^\dagger}^A(\alpha) = 0$, but $B_{x^\dagger}^A(\alpha) \not\to 0$ as $\alpha \to 0$ when the solution $x^\dagger$ to (1.12) fails to satisfy (3.6).

The following considerations are only of interest for the ill-posed case, because the estimate (3.8) below is not helpful for the well-posed case, in which (2.12) directly yields the linear rate (2.7) for all $x^\dagger \in X$. In the ill-posed case, however, the asymptotics of $B^A_{x^\dagger}(\alpha)$ for $\alpha \to 0$ determines, for example by equilibrating the two terms in the right hand side of the inequality

$$\|x^\delta_\alpha - x^\dagger\| \le \alpha \, \|(A+\alpha I)^{-1}(x^\dagger - \bar{x})\| + \frac{\delta}{\alpha}, \tag{3.8}$$

the chances and limitations of possible convergence rates of the total regularization error. This point was intensively analyzed in [40] for fractional power source conditions

$$x^\dagger - \bar{x} = A^p \, w, \qquad w \in X, \quad 0 < p \le 1, \tag{3.9}$$

yielding for all $0 < p \le 1$ the Hölder type convergence rates of the bias

$$B^A_{x^\dagger}(\alpha) = O(\alpha^p) \qquad \text{as} \quad \alpha \to 0. \tag{3.10}$$

We note that for monotone operators $A \in \mathcal{L}(X)$ the fractional powers A^p, $0 < p \le 1$, are defined via the Balakrishnan calculus as

$$A^p := \frac{\sin \pi p}{\pi} \int_0^\infty s^{p-1} \, (A + sI)^{-1} \, A \; ds. \tag{3.11}$$

Because of the saturation result for the bias presented with the following theorem, we call the source condition

$$x^\dagger - \bar{x} = A \, w, \quad w \in X, \tag{3.12}$$

benchmark source condition.

Theorem 3.4. *With the exception of the singular case $x^\dagger - \bar{x} = 0$ the benchmark source condition (3.12) yields with*

$$B^A_{x^\dagger}(\alpha) = O(\alpha) \qquad \textit{as} \quad \alpha \to 0 \tag{3.13}$$

the best possible bias rate, because

$$B^A_{x^\dagger}(\alpha) = o(\alpha) \qquad \textit{as} \quad \alpha \to 0 \qquad \textit{implies} \qquad x^\dagger - \bar{x} = 0. \tag{3.14}$$

Proof. Under the benchmark source condition we have $B^A_{x^\dagger}(\alpha) = \alpha \|(A+\alpha I)^{-1} A w\| \le \alpha \|w\| = O(\alpha)$ as $\alpha \to 0$ due to (1.14). To prove the implication (3.14) we distinguish the cases $A = 0$ and $A \ne 0$. For $A = 0$ we have $B^0_{x^\dagger}(\alpha) = \|x^\dagger - \bar{x}\|$, and the implication (3.14) is evidently true. In the case $A \ne 0$ we conclude as follows: For all $x \in X$ it holds $\|(A+\alpha I)x\| \le \|Ax\| + \alpha\|x\| \le (\|A\| + \alpha)\|x\|$. Moreover, we have $\|x^\dagger - \bar{x}\| \le (\|A\| + \alpha)\|(A+\alpha I)^{-1}(x^\dagger - \bar{x})\|$ for arbitrary $x^\dagger - \bar{x} \in X$, which is for all $\alpha > 0$ equivalent to

$$\frac{\|x^\dagger - \bar{x}\|}{\|A\| + \alpha} \;\le\; \|(A+\alpha I)^{-1}(x^\dagger - \bar{x})\|. \tag{3.15}$$

On the other hand, from (3.15) we derive for $A \ne 0$ the inequality

$$\liminf_{\alpha \to 0} \|(A+\alpha I)^{-1}(x^\dagger - \bar{x})\| \;\ge\; \frac{\|x^\dagger - \bar{x}\|}{\|A\|},$$

which for $x^\dagger - \bar{x} \neq 0$ and $\frac{\|x^\dagger - \bar{x}\|}{\|A\|} > 0$ violates the limit condition

$$\lim_{\alpha \to 0} \frac{B^A_{x^\dagger}(\alpha)}{\alpha} = \lim_{\alpha \to 0} \|(A + \alpha I)^{-1}(x^\dagger - \bar{x})\| = 0.$$

This completes the proof. □

As already mentioned, the rates of the bias $B^A_{x^\dagger}(\alpha)$ for $\alpha \to 0$ also determine the total error profile on the basis of the estimate (3.8). In the simplest case of an a priori choice $\alpha = \alpha(\delta) \sim \delta^{\frac{1}{p+1}}$ we directly derive for all $0 < p \leq 1$ Hölder convergence rates

$$\|x^\delta_{\alpha(\delta)} - x^\dagger\| = O(\delta^{\frac{p}{p+1}}) \qquad \text{as} \quad \delta \to 0 \tag{3.16}$$

from the source conditions (3.9). In the benchmark case $p = 1$ this gives

$$\|x^\delta_{\alpha(\delta)} - x^\dagger\| = O(\delta^{\frac{1}{2}}) \qquad \text{as} \quad \delta \to 0, \tag{3.17}$$

and Plato's saturation theorem [35, Theorem 5.1] (cf. formula (2.11) of Proposition 2.6) proves that for the ill-posed situation this is the maximal best possible error rate for the Lavrentiev regularization in the linear case, since $\|x^\delta_{\alpha(\delta, y^\delta)} - x^\dagger\| = o(\delta^{\frac{1}{2}})$ implies for arbitrary a posteriori choices $\alpha = \alpha(\delta, y^\delta)$ of the regularization parameter that $x^\dagger - \bar{x} = 0$.

Recently, it was shown in [36] that alternative source conditions

$$x^\dagger - \bar{x} = (A^*)^p\, w, \quad w \in X,$$

which replace the monotone non-selfadjoint operator A with the also monotone adjoint A^*, can be less efficient with respect to rate results if $1/2 \leq p \leq 1$. In the worst case, the best possible bias rate under the adjoint source condition $x^\dagger - \bar{x} = A^* w$ is $B^A_{x^\dagger}(\alpha) = O(\sqrt{\alpha})$. This worst case, for example, takes place with $X := L^2(0,1)$ when we consider the Riemann-Liouville fractional integral operator $A := V$ studied below in Example 4.7. This case is connected with a reduced total error rate $\|x^\delta_\alpha - x^\dagger\| = O(\delta^{\frac{1}{3}})$ for $p = 1$ in comparison to (3.17). Consequently, the situation of Lavrentiev regularization differs significantly from the situation of Tikhonov's regularization method, where just this adjoint source condition is advantageous (cf. [17, Corollary 3.1.3]).

In Section 4, by exploiting the above mentioned bias studies and by using the method of approximate source conditions with benchmark condition (3.12), we will extend the results to general, non-Hölder type, and low order convergence rates occurring in the context of linear Lavrentiev regularization. We note that the focus is on non-selfadjoint operators A, where spectral theory fails. Since the solution-independent bound (3.4) for the noise propagation error is also valid for the Lavrentiev regularization (1.9) applied to nonlinear equations (1.1), we will show in Section 5 that a bias-based error analysis can also be successful for classes of monotone forward operators F under specific restrictions of the nonlinearity structure.

4 General convergence rates for the linear case using approximate source conditions

Based on the paper [31], for a given selfadjoint non-negative linear operator $H \in \mathcal{L}(X)$ with non-closed range $\mathcal{R}(H) \neq \overline{\mathcal{R}(H)}$ in the Hilbert space X, it can be shown that for

every element $u \in X$ with $u \perp \mathcal{N}(H)$ there exist an index function φ and a source element $w \in X$ such that a general source condition $u = \varphi(H)w$ holds, where $\varphi(H)$ as usual is defined by spectral theory such that any spectrum point $\lambda > 0$ of H corresponds to the spectrum point $\varphi(\lambda)$ of $\varphi(H)$. For a selfadjoint monotone operator $A \in \mathcal{L}(X)$ with non-closed range this proves with $H := A$ under (3.6) a source condition

$$x^\dagger - \bar{x} = \varphi(A)w, \quad w \in X. \tag{4.1}$$

Using spectral properties of A this makes it possible to formulate convergence rates $B^A_{x^\dagger}(\alpha) = O(\psi(\alpha))$ as $\alpha \to 0$ for the Lavrentiev regularization bias with some index function ψ which depends on the index function φ, as is similarly done for the Tikhonov regularization bias with $H := A^*A$ (cf., e.g., [2, 14, 31]). Some authors exploit this approach for the Lavrentiev regularization, partially even in a nonlinear setting, see [4, 30, 32, 39], but their restriction to selfadjoint monotone operators $A \in \mathcal{L}(X)$ is rather artificial, because in particular the case of non-selfadjoint monotone linear operators (see Examples 4.7 and 4.8 below) is of interest. For such operators A, however, spectral theory is not applicable and the Balakrishnan calculus (cf. (3.11)) only allows us to handle power type source conditions (3.9) yielding Hölder convergence rates (3.10) for the bias $B^A_{x^\dagger}(\alpha)$ as $\alpha \to 0$ and consequently yielding only Hölder convergence rates $\|x^\delta_\alpha - x^\dagger\| = O(\delta^{\frac{p}{p+1}})$ as $\delta \to 0$ for $0 < p \le 1$ and noisy data when taking into account the estimates (3.3) and (3.4).

As already mentioned in [23, Section 4.1], by avoiding expressions $\varphi(A)$ with index functions φ of non-power type and non-selfadjoint monotone linear operators A, the method of approximate source conditions can help to verify low order convergence rates of non-Hölder type for the Lavrentiev regularization without self-adjontness assumptions of the forward operator in the linear case or of its Fréchet derivative at the solution in the nonlinear case. We will outline details of such an approach for the linear case in this section and for the nonlinear case in the subsequent section. The method of approximate source conditions had been developed for linear ill-posed operator equations in Hilbert spaces in [21] (see also [12]) and was extended to nonlinear equations and a Banach space setting in [19] and [7] (see also [38]). Associated with the best possible rate (3.13) the condition (3.12) acts in an optimal manner as benchmark source condition for obtaining convergence rates if $x^\dagger - \bar{x}$ satisfies (3.6), but violates (3.12). In such case the smoothness of the element $x^\dagger - \bar{x}$ with respect to the monotone operator A is too small for having the bias rate (3.13) and one can use the distance function

$$d(R) = \min\{\|x^\dagger - \bar{x} - Aw\| : \|w\| \le R\}, \qquad 0 \le R < \infty, \tag{4.2}$$

to measure for $x^\dagger - \bar{x}$ the degree of violation with respect to the benchmark condition expressed by the decay rate of $d(R) \to 0$ as $R \to \infty$.

Lemma 4.1. *Assume that for the monotone operator $A \in \mathcal{L}(X)$ the element $x^\dagger \in X$ fails the benchmark source condition (3.12), i.e. $x^\dagger - \bar{x} \notin \mathcal{R}(A)$, but satisfies the orthogonality condition (3.6). Then the distance function $d(R)$ from (4.2) is positive, strictly decreasing and concave and hence continuous for all $0 \le R < \infty$ and satisfies the limit condition $\lim_{R\to\infty} d(R) = 0$.*

Proof. The assertion of the lemma follows immediately from [7, Lemma 3.2] if the condition $x^\dagger - \bar{x} \in \overline{\mathcal{R}(A)}$ is valid. Since $x^\dagger - \bar{x} \perp \mathcal{N}(A)$ implies that $x^\dagger - \bar{x} \in \overline{\mathcal{R}(A^*)}$

(see, e.g., [5, Fact 2.18 (iii)]), this condition however is a consequence of the identity $\overline{\mathcal{R}(A)} = \overline{\mathcal{R}(A^*)}$ for monotone operators $A \in \mathcal{L}(X)$ (see, e.g., [5, Proposition 20.17]). □

We easily derive that for $x^\dagger - \bar{x}$ from Lemma 4.1 and arbitrary $R > 0$ there exist elements $w_R \in X$ with $\|w_R\| = R$ and $r_R \in X$ with $\|r_R\| = d(R)$ such that an approximate source condition of the form

$$x^\dagger - \bar{x} = Aw_R + r_R \tag{4.3}$$

is valid. Then due to (1.14) we have

$$B^A_{x^\dagger}(\alpha) = \alpha\|(A+\alpha I)^{-1}(Aw_R + r_R)\|$$

$$\leq \alpha\|(A+\alpha I)^{-1}A\|\|w_R\| + \alpha\|(A+\alpha I)^{-1}\|\|r_R\| \leq R\alpha + d(R),$$

and equilibrating the last two terms by means of the strictly decreasing auxiliary function

$$\Phi(R) := \frac{d(R)}{R}, \qquad 0 < R < \infty, \qquad \lim_{R\to\infty} d(R) = 0, \tag{4.4}$$

as $R := \Phi^{-1}(\alpha)$ we have the assertion of the following proposition.

Proposition 4.2. *Under the assumptions of Lemma 4.1 we have the bias estimate*

$$B^A_{x^\dagger}(\alpha) \leq 2d(\Phi^{-1}(\alpha)), \qquad \alpha > 0, \tag{4.5}$$

for the Lavrentiev regularization in the linear case, where $d(\Phi^{-1}(\alpha))$ is an index function with $\lim_{\alpha\to 0} \frac{\alpha}{d(\Phi^{-1}(\alpha))} = 0$. Moreover, for the a priori parameter choice $\alpha(\delta) := \Psi^{-1}(\delta)$ with $\Psi(\alpha) := \alpha\, d(\Phi^{-1}(\alpha))$ we have from (3.3) and (3.4) the estimate

$$\|x^\delta_{\alpha(\delta)} - x^\dagger\| \leq 3d(\Phi^{-1}(\Psi^{-1}(\delta))), \qquad \delta > 0,$$

for the total regularization error and hence the convergence rate

$$\|x^\delta_{\alpha(\delta)} - x^\dagger\| = O(d(\Phi^{-1}(\Psi^{-1}(\delta)))) \qquad \text{as} \quad \delta \to 0.$$

We note that the assertion of Proposition 4.2 remains valid if $d(R)$ beginning with (4.4) is replaced with a concave majorant of the distance function.

Special case 4.3. (Distance functions with power-type decay rate) As a consequence of the range identity

$$\mathcal{R}(A^p) = \mathcal{R}((AA^*)^{p/2}),$$

proven for all $0 < p \leq 1$ in [36, Lemma 1], and on the basis of the assertion in [12, Theorem 3.2] we have that for all $0 < p < 1$ the fractional power source conditions (3.9) lead to distance functions (cf. (4.2)) with a power-type decay as

$$d(R) \leq C\, R^{\frac{p}{p-1}}, \qquad C > 0. \tag{4.6}$$

The smaller $p > 0$ the slower is the decay rate of $d(R) \to 0$ as $R \to \infty$ and the higher is for $x^\dagger - \bar{x}$ the degree of violation with respect to the benchmark source condition (3.12). Applying Proposition 4.2 this yields with $\Phi(R) \sim R^{\frac{1}{p-1}}$, $\Psi(\alpha) \sim \alpha^{p+1}$ and $\Psi^{-1}(\delta) \sim \delta^{\frac{1}{p+1}}$ the Lavrentiev regularization error convergence rates

$$B^A_{x^\dagger}(\alpha) = O(\alpha^p) \qquad \text{and} \qquad \|x^\delta_{\alpha(\delta)} - x^\dagger\| = O(\delta^{\frac{p}{p+1}}) \quad \text{if} \quad \alpha(\delta) \sim \delta^{\frac{1}{p+1}}.$$

Special case 4.4. (Distance functions with logarithmic decay rate) If for $x^\dagger - \bar{x}$ the degree of violation with respect to the benchmark source condition (3.12) is so extreme that the power type decay (4.6) cannot hold for arbitrarily small $p > 0$, then a very slow logarithmic decay rate

$$d(R) \leq \frac{K}{(\log R)^q}, \qquad K, q > 0, \tag{4.7}$$

is still possible for sufficiently large $R \geq \underline{R} > 0$. Then the derived formula

$$B^A_{x^\dagger}(\alpha) \leq R\alpha + d(R) \tag{4.8}$$

applies by setting $R := \alpha^{-\kappa}$, $0 < \kappa < 1$, and provides us with the estimate

$$B^A_{x^\dagger}(\alpha) \leq \alpha^{1-\kappa} + \frac{K}{[\log(\alpha^{-\kappa})]^q} = \alpha^{1-\kappa} + \frac{K}{\kappa^q \left[\log\left(\frac{1}{\alpha}\right)\right]^q}$$

and owing to $\alpha^{1-\kappa} = O\left(\left[\log\left(\frac{1}{\alpha}\right)\right]^q\right)$ as $\alpha \to 0$ we have

$$B^A_{x^\dagger}(\alpha) = O\left(\left[\log\left(\frac{1}{\alpha}\right)\right]^q\right) \quad \text{as} \quad \alpha \to 0. \tag{4.9}$$

This logarithmic convergence rate for the bias also leads to a logarithmic rate for the noisy data case of linear Lavrentiev regularization. Namely, we derive from (4.9) in combination with (3.3) and (3.4) the convergence rate

$$\|x^\delta_{\alpha(\delta)} - x^\dagger\| = O\left(\left[\log\left(\frac{1}{\delta}\right)\right]^q\right) \quad \text{as} \quad \delta \to 0 \tag{4.10}$$

when the regularization parameter is chosen a priori as $\alpha(\delta) \sim \delta^\zeta$ with exponent $0 < \zeta < 1$. The very low logarithmic convergence rates of the form (4.10) are well-known in regularization theory (see, e.g, [7, 20, 27, 41]) and inevitable if the solution is not smooth enough with respect to the forward operator.

Remark 4.5. Under the conditions $x^\dagger - \bar{x} \notin \mathcal{R}(A)$ and $x^\dagger - \bar{x} \perp \mathcal{N}(A)$ the properties of the distance function $d(R)$ from (4.2) are such that the inequality (4.8) is valid for sufficiently large $R \geq \underline{R} > 0$. Then the function $\Phi(R)$ defined by formula (4.4) is strictly decreasing for those R and tends to zero as $R \to \infty$. Moreover, the mapping $R \mapsto \alpha := \Phi(R)$ is injective (strictly decreasing) and well-defined for $\underline{R} \leq R < \infty$ with $\alpha \to 0$ as $R \to \infty$. Then it is clear that the function $d(\Phi^{-1}(\alpha))$ is well-defined and strictly increasing for sufficiently small $\alpha > 0$ with the limit condition $\lim\limits_{\alpha \to 0} d(\Phi^{-1}(\alpha)) = 0$. This, however yields the bias convergence

$$\lim_{\alpha \to 0} B^A_{x^\dagger}(\alpha) = 0$$

under the above stated conditions on $x^\dagger - \bar{x}$.

Example 4.6. (Multiplication operator) We start this list of examples with the simple multiplication operator $A := M$ in the real Hilbert space $X := L^2(0,1)$ generated by a real multiplier function $m \in L^\infty(0,1)$, which is defined as

$$[Mx](t) := m(t)\, x(t), \qquad 0 \leq t \leq 1, \tag{4.11}$$

where we restrict to continuous and strictly increasing functions m with $\lim_{t\to 0} m(t) = 0$ and $\lim_{t\to 1} m(t) = 1$. Then the linear operator M is bounded, injective, non-compact with continuous spectrum $\sigma(M) = [0,1]$, selfadjoint and monotone, i.e. the associated linear operator equation (1.12) is ill-posed. The smoothness of the solution $x^\dagger$ with respect to M and $\bar{x}$ and the decay rate of the distance function $d(R)$ if the benchmark source condition is violated, i.e. if $(x^\dagger - \bar{x})/m \notin L^2(0,1)$, will be determined by the decay rate of $m(t) \to 0$ as $t \to 0$. For example, in the situation $x^\dagger - \bar{x} \equiv 1$ we have a logarithmic decay of the distance function $d(R) \leq K/\log(R)$ for sufficiently large R and hence logarithmic convergence rates (4.10) with $q = 1$ for $m(t) = \exp(1 - 1/\sqrt{t})$ (see for details [14, Example 3]).

Example 4.7. (Fractional integral operator) As second point in this series we present in the real Hilbert space $X := L^2(0,1)$ the Riemann-Liouville fractional integral operator (cf. [15]) $A := V$, also called Volterra operator (cf. [18]), defined by

$$[Vx](s) := \int_0^s x(t)dt, \qquad 0 \leq s \leq 1. \tag{4.12}$$

This is an example which helps to distinguish different situations of solution smoothness with respect to the forward operator. The linear operator V is bounded, injective, compact, non-selfadjoint and monotone with $0 \in \sigma(V)$, i.e. the associated linear operator equation (1.12) is ill-posed. On the hand, we have situations, where $x^\dagger$ allows for Hölder source conditions (3.9) implying Hölder rates (3.10) for the bias of the Lavrentiev regularization. On the other hand, there exist elements $x^\dagger$ with even less smoothness such that (3.9) fails for arbitrarily small exponents $p > 0$. Then only lower rates, for example logarithmic rates (4.9), remain for the bias taking into account (cf. Remark 3.3) that we have for all $x^\dagger \in X$ with $x^\dagger - \bar{x} \perp \mathcal{N}(A)$ the limit condition $B^A_{x^\dagger}(\alpha) \to 0$ as $\alpha \to 0$ and hence a (perhaps very low) well-defined bias rate for that specific $x^\dagger$.

From [15, Theorem 2.1] (see also [16]) we take the explicit structure of the ranges of the fractional powers of the integral operator (4.12) in terms of fractional order Sobolev-Hilbert spaces $H^p(0,1)$ as

$$\mathcal{R}(V^p) = \begin{cases} H^p[0,1] & \text{for} \quad 0 < p < \frac{1}{2} \\ \{u \in H^{\frac{1}{2}}[0,1] : \int_0^1 \frac{|u(t)|^2}{t} dt < \infty\} & \text{for} \quad p = \frac{1}{2} \\ \{u \in H^p[0,1] : \, u(0) = 0\} & \text{for} \quad \frac{1}{2} < p \leq 1 \end{cases} \tag{4.13}$$

Moreover, due to the injectivity of the Volterra operator V in $L^2(0,1)$, the orthogonality condition $x^\dagger - \bar{x} \perp \mathcal{N}(A)$ and hence the bias limit $B^V_{x^\dagger}(\alpha) \to 0$ as $\alpha \to 0$ are trivially satisfied. Now, for all $0 < p < 1/2$, Hölder source conditions (3.9) hold true for $A := V$ if and only if $x^\dagger - \bar{x} \in H^p(0,1)$. For $1/2 \leq p \leq 1$ additional conditions on $x^\dagger - \bar{x}$ have to be imposed. Then, with that non-selfadjoint monotone forward operator V, a necessary condition for the situation that a logarithmic bias rate (4.9) is valid, but (3.10) fails for arbitrarily small $p > 0$, is

$$x^\dagger, \bar{x} \in L^2(0,1), \quad \text{but} \quad x^\dagger - \bar{x} \notin H^p(0,1) \quad \text{for arbitrarily small} \quad p > 0.$$

Example 4.8. (Cesàro operator) As third example we consider in $X := L^2(0,1)$ the continuous version of the Cesàro operator $A := C$ (cf. [8, p. 133]) defined as

$$[Cx](s) := \frac{1}{s}\int_0^s x(t)dt, \qquad 0 \le s \le 1, \tag{4.14}$$

which is, in contrast to (4.12), an injective, monotone, but non-compact linear operator. Again we have $0 \in \sigma(C)$ and hence ill-posedness, but the conditions $x^\dagger - \bar{x} \in \mathcal{R}(C^p)$ for exponents $0 < p \le 1$ are not so easy to check like in Example 4.7 in order to derive Hölder convergence rates (3.16). This is a good example for proving the capability of distance functions for that purpose. Evidently, $x^\dagger - \bar{x} \equiv 1$ satisfies the benchmark source condition (3.12) with source element $w \equiv 1$, but the Heaviside-type function

$$x^\dagger(t) - \bar{x}(t) = \begin{cases} 0 & \text{if} \quad 0 \le t < 1/2 \\ 1 & \text{if} \quad 1/2 \le t \le 1 \end{cases} \tag{4.15}$$

fails to satisfy (3.12) for arbitrary $w \in L^2(0,1)$. However, as Proposition 4.9 will show, we have $d(R) \le \frac{K}{R}$ for some constant $K > 0$ and sufficiently large $R > 0$, which yields the inequality (4.6) with $p = 1/2$ and hence the Hölder rate $\mathcal{O}(\delta^{1/3})$ for the Lavrentiev regularization.

Proposition 4.9. *For $x^\dagger - \bar{x}$ from (4.15) we have for some constant $K > 0$ and sufficiently large $R > 0$ the estimate*

$$d(R) \;=\; \min\{\|x^\dagger - \bar{x} - Cw\| : \|w\| \le R\} \;\le\; \frac{K}{R}\,.$$

Proof. For the function $w_R(t) = \begin{cases} 0 & \text{if} \quad 0 \le t < 1/2 \\ R^2 & \text{if} \quad 1/2 \le t \le 1/2 + 1/(2R^2) \\ 1 & \text{if} \quad 1/2 + 1/(2R^2) < t \le 1 \end{cases}$ we have

$$\|w_R\| \le R \quad \text{and} \quad d^2(R) \le \int_0^1 \left(\frac{1}{s}\int_0^s w_R(t)dt - (x^\dagger(s) - \bar{x}(s)) \right)^2 ds\,.$$

Furthermore, because w_R and $x^\dagger - \bar{x}$ are identically zero on $[0, 1/2)$ we can estimate as

$$\begin{aligned} d^2(R) &\le \int_0^1 \frac{1}{s^2} \left(\int_0^s w_R(t)dt - s(x^\dagger(s) - \bar{x}(s)) \right)^2 ds \\ &\le 4 \int_{1/2}^1 \left(\int_{1/2}^s w_R(t)dt - s(x^\dagger(s) - \bar{x}(s)) \right)^2 ds \\ &= 4 \left[\int_0^{1/(2R^2)} \left(s(R^2-1) - \frac{1}{2} \right)^2 ds + \frac{1}{4R^4} \left(\frac{1}{2} - \frac{1}{2R^2} \right) \right] \\ &= 4 \left[\frac{(R^2-1)^2}{24R^6} - \frac{R^2-1}{8R^4} + \frac{1}{8R^2} + \frac{1}{8R^4} - \frac{1}{8R^6} \right] \le K^2 \frac{1}{R^2} \end{aligned}$$

for sufficiently large $R > 0$. This completes the proof. □

5 Extensions to the nonlinear case under specific conditions on the nonlinearity structure

Now we return to the Lavrentiev regularization for nonlinear operator equations (1.1) with regularized solutions x_α^δ satisfying for noisy data y^δ the singularly perturbed equation (1.9) and with regularized solutions x_α satisfying (3.1) in the noise-free case. We are going to handle the corresponding nonlinear bias $B_{x^\dagger}^F(\alpha) := \|x_\alpha - x^\dagger\|$ in the noise-free case, where we try to incorporate experiences from the studies of the linear case in Section 4. By Lemma 3.1 the properties of the bias $B_{x^\dagger}^F(\alpha)$ allow us immediately to derive the essential behaviour of the total error $\|x_\alpha^\delta - x^\dagger\|$ for the Lavrentiev regularization also in the nonlinear case.

Throughout this section let us suppose that the following assumption holds in addition to Assumption 1.2.

Assumption 5.1.

(i) Let there exist a ball $\mathcal{B}_r(x^\dagger) \subset X$ around a solution $x^\dagger$ to equation (1.1) with sufficiently large radius

$$r > \|x^\dagger - \bar{x}\| \tag{5.1}$$

such that F is Fréchet differentiable in the ball with Fréchet derivatives $F'(x) \in \mathcal{L}(X)$ and the mapping $x \mapsto F'(x)$ is continuous at every $x \in \mathcal{B}_r(x^\dagger)$.

(ii) Let there exist a constant $k_0 > 0$ and a function g such that, for every $\tilde{x}, x \in \mathcal{B}_r(x^\dagger)$ and $v \in X$, there is $g(\tilde{x}, x, v) \in X$ satisfying the nonlinearity condition

$$(F'(\tilde{x}) - F'(x))\, v = F'(x)\, g(\tilde{x}, x, v), \qquad \|g(\tilde{x}, x, v)\| \le k_0 \|\tilde{x} - x\|\, \|v\|. \tag{5.2}$$

Item (ii) of Assumption 5.1 occurs in numerous papers on regularization theory in a more or less modified form, but we follow the ideas of Mahale and Nair presented in [30] which have filled gaps of the previous literature (see for details [30, p. 195])). Furthermore, we mention at this point that example classes of nonlinear forward operators F satisfying the specific nonlinearity condition (5.2) were presented, for example, in [37] and [3, 9].

Since 2000, in a series of papers by Tautenhahn, Nair, Mahale, George and others, partially written with coauthors, order optimal convergence rates for the Lavrentiev regularization have been proven for linear as well as for nonlinear equations with monotone forward operators under a priori and a posteriori choices of the regularization parameter $\alpha > 0$. The focus was always on Hölder rates under Hölder-type source conditions. The papers of these authors on general source conditions exploiting terms of the form $\varphi(A)$ or $\varphi(F'(x^\dagger))$ with index functions φ of non-power type suffer from the fact that such terms are only well-defined for selfadjoint monotone operators A and $F'(x^\dagger)$. This drawback is now compensated by our approximate source condition approach using distance functions.

Now we are ready to formulate the proposition of this section, which assert that under the assumed nonlinearity conditions and for a sufficiently good reference element $\bar{x}$ the bias $B_{x^\dagger}^F(\alpha)$ in the nonlinear case is proportional to and hence fully determined by the bias function $B_{x^\dagger}^A(\alpha)$, where A is the monotone operator $F'(x^\dagger) \in \mathcal{L}(X)$. Thus, convergence as well as Hölder or logarithmic rates for the noise-free bias $B_{x^\dagger}^A(\alpha)$ as $\alpha \to 0$ carry over to

the same rates for the bias $B^F_{x^\dagger}(\alpha)$ with corresponding consequences for the convergence and for the rates of the total regularization errors $\|x^\delta_\alpha - x^\dagger\|$ as $\delta \to 0$.

Theorem 5.2. *Under Assumptions 1.2 and 5.1 with the additional condition $k_0 \|x^\dagger - \bar{x}\| < 2$ let the operator $A := F'(x^\dagger)$ denote the Fréchet derivative of the monotone nonlinear operator F at the solution $x^\dagger$ to equation (1.1). Then we have, for all $\alpha > 0$ and associated elements x_α solving (3.1), the inequality*

$$B^F_{x^\dagger}(\alpha) \le \|x^\dagger - \bar{x}\| \tag{5.3}$$

and thus from (5.1) the condition $x_\alpha \in \mathcal{B}_r(x^\dagger)$. Furthermore, we have the estimate

$$B^F_{x^\dagger}(\alpha) \le C\, B^A_{x^\dagger}(\alpha), \qquad \text{with} \qquad C = \frac{2 + 2k_0 \|x^\dagger - \bar{x}\|}{2 - k_0 \|x^\dagger - \bar{x}\|}\,. \tag{5.4}$$

Hence, we have for all $\alpha > 0$ and $\delta \ge 0$ the estimate

$$\|x^\delta_\alpha - x^\dagger\| \le C\, B^A_{x^\dagger}(\alpha) + \frac{\delta}{\alpha} \tag{5.5}$$

for the total regularization error.

Proof. First we quote as result from [40, Proposition 3.1] that the inequality (5.3) and by (5.1) also $x_\alpha \in \mathcal{B}_r(x^\dagger)$ are valid for all $\alpha > 0$. It is well-known that the Fréchet derivatives $A := F'(x^\dagger)$ and $A_\alpha := F'(x_\alpha)$ are monotone bounded linear operators mapping in X if the nonlinear operator F is monotone. Now, amending some ideas along the lines of the proofs of [30, Proposition 2.2] and [40, Proposition 3.3] we derive from (3.1) that

$$x_\alpha - x^\dagger = \bar{x} - x^\dagger + (A_\alpha + \alpha I)^{-1}[y - F(x_\alpha) + A_\alpha(x_\alpha - \bar{x})] = u_\alpha + v_\alpha + w_\alpha,$$

with the three terms

$$u_\alpha := \alpha(A + \alpha I)^{-1}(\bar{x} - x^\dagger), \qquad \text{where} \qquad \|u_\alpha\| = B^A_{x^\dagger}(\alpha),$$

$$v_\alpha := (A_\alpha + \alpha I)^{-1}(A - A_\alpha)\, u_\alpha\,,$$

and

$$w_\alpha := (A_\alpha + \alpha I)^{-1}[F(x^\dagger) - F(x_\alpha) + A_\alpha(x_\alpha - x^\dagger)]\,.$$

The handling of the term v_α is based on the nonlinearity condition (5.2), where we have with $(A - A_\alpha)\, u_\alpha = A_\alpha\, g(x^\dagger, x_\alpha, u_\alpha)$ the estimate

$$\|v_\alpha\| = \|(A_\alpha + \alpha I)^{-1} A_\alpha\, g(x^\dagger, x_\alpha, u_\alpha\| \le k_0 \|x_\alpha - x^\dagger\| \|u_\alpha\|\,.$$

Owing to the continuous Fréchet differentiability of F in the ball $\mathcal{B}_r(x^\dagger)$ the fundamental theorem of calculus (mean value theorem in integral form) applies to estimate from above the norm $\|w_\alpha\|$ of the third term. Precisely, we have

$$F(x^\dagger) - F(x_\alpha) - A_\alpha(x^\dagger - x_\alpha) = \int_0^1 [F'(x_\alpha + t(x^\dagger - x_\alpha)) - A_\alpha](x^\dagger - x_\alpha)\, dt$$

$$= A_\alpha \int_0^1 g(x_\alpha + t(x^\dagger - x_\alpha), x_\alpha, x^\dagger - x_\alpha)\, dt$$

and

$$\|g(x_\alpha + t(x^\dagger - x_\alpha), x_\alpha, x^\dagger - x_\alpha)\| \le k_0 \|x_\alpha - x^\dagger\|^2 t\,.$$

Then we have due to the second inequality in (1.14)

$$\|w_\alpha\| = \|(A_\alpha + \alpha I)^{-1} A_\alpha \int_0^1 g(x_\alpha + t(x^\dagger - x_\alpha), x_\alpha, x^\dagger - x_\alpha)\, dt\| \le \frac{k_0}{2} \|x_\alpha - x^\dagger\|^2\,.$$

Summarizing the results for u_α, v_α and w_α we obtain

$$\|x_\alpha - x^\dagger\| \le \left(1 + k_0 \|x_\alpha - x^\dagger\|\right) \|u_\alpha\| + \frac{k_0}{2} \|x_\alpha - x^\dagger\|^2$$

and taking into account (5.3)

$$\|x_\alpha - x^\dagger\| \le \left(1 + k_0 \|x^\dagger - \bar{x}\|\right) \|u_\alpha\| + \frac{k_0}{2} \|x^\dagger - \bar{x}\| \|x_\alpha - x^\dagger\|\,.$$

This yields (5.4) as

$$B^F_{x^\dagger}(\alpha) = \|x_\alpha - x^\dagger\| \le \left(\frac{2 + 2k_0 \|x^\dagger - \bar{x}\|}{2 - k_0 \|x^\dagger - \bar{x}\|}\right) \|u_\alpha\| = \left(\frac{2 + 2k_0 \|x^\dagger - \bar{x}\|}{2 - k_0 \|x^\dagger - \bar{x}\|}\right) B^A_{x^\dagger}(\alpha).$$

The estimate (5.5) is a consequence of Lemma 3.1 (cf. formula (3.4)) in combination with the triangle inequality (3.3). This completes the proof. □

The following corollary from Theorem 5.2 is a counterpart to Proposition 3.2 concerning the convergence of the bias as the regularization parameter α tends to zero.

Corollary 5.3. *Under the assumptions of Theorem 5.2 we have*

$$\lim_{\alpha \to 0} B^F_{x^\dagger}(\alpha) = 0$$

if

$$x^\dagger - \bar{x} \perp \mathcal{N}(F'(x^\dagger)).$$

Remark 5.4. The very specific nonlinearity condition (5.2) can be avoided if the benchmark source condition in the nonlinear case

$$x^\dagger - \bar{x} = F'(x^\dagger)\, w, \quad w \in X,$$

applies together with the simpler Lipschitz continuity

$$\|F'(x) - F'(x^\dagger)\| \le L \|x - x^\dagger\| \qquad \text{for all} \qquad x \in \mathcal{B}_r(x^\dagger)$$

as nonlinearity condition for some $L > 0$. Then we have from [40, Theorem 3.2] the nonlinear bias estimate

$$B^F_{x^\dagger}(\alpha) \le \left(\|w\| + \frac{L}{2}\|w\|^2\right) \alpha,$$

which yields together with (3.3) and (3.4) the convergence rate

$$\|x^\delta_{\alpha(\delta)} - x^\dagger\| = \mathcal{O}(\sqrt{\delta}) \qquad \text{as} \qquad \delta \to 0$$

for the total regularization error if the regularization parameter is chosen as $\alpha(\delta) \sim \sqrt{\delta}$.

A convergence rate result similar to that of Proposition 5.2 for the Lavrentiev regularization of nonlinear operator equations was also presented as Theorem 8 in [23]. In contrast to item (ii) of our Assumption 5.1 a range invariance occurs there as structural condition of nonlinearity for the forward operator F, which provides the opportunity to use a monotone operator $A \in \mathcal{L}(X)$ different from the Fréchet derivative $F'(x^\dagger)$. However, the cross connections between the linear bias $B^A_{x^\dagger}(\alpha)$ and its nonlinear counterpart $B^F_{x^\dagger}(\alpha)$ are not so clear in [23] as they are in Proposition 5.2.

Acknowledgments

The authors are grateful to Robert Plato (Siegen) for valuable hints and private communications. Research of BH was partially supported by the German Research Foundation (DFG) under grant HO 1454/8-2.

References

[1] Y. Alber and I. Ryazantseva: *Nonlinear Ill-posed Problems of Monotone Type*, Springer, Dordrecht, 2006.

[2] R. Andreev, P. Elbau, M. V. de Hoop, L. Qiu, and O. Scherzer: Generalized convergence rates results for linear inverse problems in Hilbert spaces. *Numer. Funct. Anal. Optim.* **36**(5), 549–566 (2015).

[3] I. K. Argyros, Y. J. Cho and S. George: Expanding the applicability of Lavrentiev regularization methods for ill-posed problems. *Bound. Value Probl.* **2013**, 114 (15pp) (2013).

[4] I. K. Argyros and S. George: Expanding the applicability of Lavrentiev regularization methods for ill-posed equations under general source condition. *Nonlinear Funct. Anal. Appl.* **19**(2), 177–192 (2014).

[5] H. H. Bauschke and P. L. Combettes: *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*, Springer, New York, 2011.

[6] A. Besenyei: On uniformly monotone operators arising in nonlinear elliptic and parabolic problems. *Ann. Univ. Sci. Budapest. Eötvös Sect. Math.* **53**, 33–43 (2010).

[7] R. I. Boţ and B. Hofmann: An extension of the variational inequality approach for obtaining convergence rates in regularization of nonlinear ill-posed problems. *Journal of Integral Equations and Applications* **22**(3), 369–392 (2010).

[8] A. Brown, A. P. R. Halmos, and A. L. Shields: Cesàro operators. *Acta Sci. Math. (Szeged)* **26**, 125–137 (1965).

[9] M. Burger and B. Kaltenbacher: Regularizing Newton-Kaczmarz methods for nonlinear ill-posed problems. *SIAM J. Numer. Anal.* **44**(1), 153–182 (2006).

[10] J. Cheng, B. Hofmann, and S. Lu: The index function and Tikhonov regularization for ill-posed problems. *J. Comput. Appl. Math.* **265**, 110–119 (2014).

[11] J. Cheng and M. Yamamoto: One new strategy for a priori choice of regularizing parameters in Tikhonov's regularization. *Inverse Problems* **16**(4), L31–L38 (2000)

[12] D. Düvelmeyer, B. Hofmann, and M. Yamamoto: Range inclusions and approximate source conditions with general benchmark functions. *Numer. Funct. Anal. Optim.* **28**(11-12), 1245–1261 (2007).

[13] H. W. Engl, M. Hanke, and A. Neubauer: *Regularization of Inverse Problems*. Kluwer Academic Publishers, Dordrecht, 1996, 2nd Edition 2000.

[14] J. Flemming, B. Hofmann, and P. Mathé: Sharp converse results for the regularization error using distance functions. *Inverse Problems*, **27**(2), 025006 (18pp) (2011).

[15] R. Gorenflo and Yu. Luchko and M. Yamamoto: Time-fractional diffusion equation in the fractional Sobolev spaces. *Fract. Calc. Appl. Anal.* **18**(3), 799–820 (2015).

[16] R. Gorenflo and M. Yamamoto: Operator-theoretic treatment of linear Abel integral equations of first kind. *Japan J. Indust. Appl. Math.* **16**(1), 137–161 (1999).

[17] C. W. Groetsch: *The Theory of Tikhonov Regularization for Fredholm Equations of the First Kind*, Pitman, Boston, MA, 1984.

[18] M. Haase: *The Functional Calculus for Sectorial Operators*, Operator Theory: Advances and Applications, Vol. 169, Birkhäuser Verlag, Basel, 2006.

[19] T. Hein and B. Hofmann: Approximate source conditions for nonlinear ill-posed problems—chances and limitations. *Inverse Problems* **25**(3), 035003 (16pp) (2009).

[20] T. Hohage and F. Weidling: Verification of a variational source condition for acoustic inverse medium scattering problems. *Inverse Problems* **31**(7), 075006 (14pp) (2015).

[21] B. Hofmann: Approximate source conditions in Tikhonov-Phillips regularization and consequences for inverse problems with multiplication operators. *Math. Methods Appl. Sci.* **29**(3), 351–371 (2006).

[22] B. Hofmann. On smoothness concepts in regularization for nonlinear inverse problems in Banach spaces. Chapter 8 in *Mathematical and Computational Modeling: With Applications in Natural and Social Sciences, Engineering, and the Arts* (Ed.: R. Melnik). John Wiley, New Jersey 2015, pp. 192–221.

[23] B. Hofmann, B. Kaltenbacher, and E. Resmerita: Lavrentiev's regularization method in Hilbert spaces revisited. *Inverse Probl. Imaging* **10**(3), 741–764 (2016).

[24] B. Hofmann and P. Mathé: Analysis of profile functions for general linear regularization methods. *SIAM J. Numer. Anal.* **45**(3), 1122–1141 (2007).

[25] B. Hofmann and O. Scherzer: Factors influencing the ill-posedness of nonlinear problems. *Inverse Problems* **10**(6), 1277–1297 (1994).

[26] B. Hofmann and M. Yamamoto: On the interplay of source conditions and variational inequalities for nonlinear ill-posed problems. *Appl. Anal.* **89**(11), 1705–1727 (2010).

[27] B. Kaltenbacher: A note on logarithmic convergence rates for nonlinear Tikhonov regularization. *J. Inv. Ill-Posed Prob.* **16**(1), 79–88 (2008).

[28] M. M. Lavrentiev: *Some Improperly Posed Problems of Mathematical Physics*. Springer, New York, 1967.

[29] F. Liu and M. Z. Nashed: Convergence of regularized solutions of nonlinear ill-posed problems with monotone operators. In: *Partial Differential Equations and Applications*, Lecture Notes in Pure and Appl. Math., Vol. 177, pp. 353–361. Dekker, New York, 1996.

[30] P. Mahale and M. T. Nair: Lavrentiev regularization of nonlinear ill-posed equations under general source condition. *J. Nonlinear Anal. Optim.* **4**(2), 193–204 (2013).

[31] P. Mathé and B. Hofmann: How general are general source conditions? *Inverse Problems* **24**(1), 015009 (5pp) (2008).

[32] M. T. Nair and U. Tautenhahn: Lavrentiev regularization for linear ill-posed problems under general source conditions. *Z. Anal. Anwendungen* **23**(1), 176–185 (2004).

[33] M. Z. Nashed: A new approach to classification and regularization of ill-posed operator equations. In: *Inverse and Ill-posed Problems* (Sankt Wolfgang, 1986), volume 4 of Notes Rep. Math. Sci. Engrg. Academic Press, Boston, MA, 1987, pp. 53–75.

[34] R. Plato: *Iterative and Other Methods for Linear Ill-Posed Equations*. Habilitation Thesis, Techn. Univ. Berlin, 1995.

[35] R. Plato: Converse results, saturation and quasi-optimality for Lavrentiev regularization of accretive problems. `https://arxiv.org/abs/1607.04879v1`, 2016.

[36] R. Plato, P. Mathé, and B. Hofmann: Optimal rates for Lavrentiev regularization with adjoint source conditions. Preprint 2016-03, Preprintreihe der Fakultät für Mathematik, TU Chemnitz, Germany, 2016. `http://nbn-resolving.de/urn:nbn:de:bsz:ch1-qucosa-199010`.

[37] O. Scherzer, H. W. Engl, and K. Kunisch: Optimal a posteriori parameter choice for Tikhonov regularization for solving nonlinear ill-posed problems. *SIAM J. Numer. Anal.* **30**(6), 1796–1838 (1993).

[38] T. Schuster, B. Kaltenbacher, B. Hofmann, and K. S. Kazimierski: *Regularization Methods in Banach Spaces*. Walter de Gruyter, Berlin/Boston, 2012.

[39] E. V. Semenova: Lavrentiev regularization and balancing principle for solving ill-posed problems with monotone operators. *Comput. Methods Appl. Math.* **10**(4), 444–454 (2010).

[40] U. Tautenhahn: On the method of Lavrentiev regularization for nonlinear ill-posed problems. *Inverse Problems* **18**(1), 191–207 (2002).

[41] F. Werner and T. Hohage: Convergence rates in expectation for Tikhonov-type regularization of inverse problems with Poisson data. *Inverse Problems* **28**(10), 104004 (15pp.) (2012).

[42] E. Zeidler: *Nonlinear Functional Analysis and its Applications - II/B: Nonlinear Monotone Operators*. Springer, 1990.