CONVERGENCE RESULTS AND LOW-ORDER RATES FOR NONLINEAR TIKHONOV REGULARIZATION WITH OVERSMOOTHING PENALTY TERM[*]

BERND HOFMANN[†] AND ROBERT PLATO[‡]

Abstract. For Tikhonov regularization of ill-posed nonlinear operator equations, convergence is studied in a Hilbert scale setting. We include the case of oversmoothing penalty terms, which means that the exact solution does not belong to the domain of definition of the considered penalty functional. In this case, we try to close a gap in the present theory, where Hölder-type convergence rates results have been proven under corresponding source conditions, but assertions on norm convergence for regularized solutions without source conditions are completely missing. A result of the present work is to provide sufficient conditions for convergence under a priori and a posteriori regularization parameter choice strategies without any additional smoothness assumption on the solution. The obtained error estimates moreover allow us to prove low-order convergence rates under associated (for example logarithmic) source conditions. Some numerical illustrations are also given.

Key words. ill-posed problem, inverse problem, Tikhonov regularization, oversmoothing penalty, a priori parameter choice strategy, discrepancy principle, logarithmic source condition

AMS subject classifications. 65J20, 65J15, 65J22, 47J06, 47J05

1. Introduction. The subject of this paper are nonlinear operator equations of the form

$$Fu = f^\dagger, \tag{1.1}$$

where $F : X \supset \mathcal{D}(F) \to Y$ is a nonlinear operator between infinite-dimensional Hilbert spaces X and Y with norms $\|\cdot\|$. We suppose that the right-hand side $f^\dagger \in Y$ is approximately given as $f^\delta \in Y$ satisfying the deterministic noise model

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

with the noise level $\delta \ge 0$. Throughout the paper, it is assumed that the considered equation (1.1) has a solution $u^\dagger \in \mathcal{D}(F)$ and is (at least locally at $u^\dagger$) ill-posed (cf. [14]).

For finding stable approximations to the solution $u^\dagger \in \mathcal{D}(F)$ of equation (1.1), we consider Tikhonov regularization, where the regularized solutions are minimizers of the extremal problem

$$T_\alpha^\delta(u) := \|Fu - f^\delta\|^2 + \alpha\|u - \overline{u}\|_1^2 \to \min \qquad \text{subject to} \quad u \in \mathcal{D}(F), \tag{1.3}$$

with a regularization parameter $\alpha > 0$. In this context, $\|\cdot\|_1$ is assumed to be a norm of a densely defined subspace X_1 of X, which is stronger than the original norm $\|\cdot\|$ in X. Throughout this paper, we suppose that the initial guess $\overline{u}$ in the penalty term of $T_\alpha^\delta(u)$ satisfies the condition

$$\overline{u} \in \mathcal{D} := \mathcal{D}(F) \cap X_1. \tag{1.4}$$

Precisely, we define the stronger norm $\|\cdot\|_1$ by a generator $B : X \supset \mathcal{D}(B) \to X$, which is a self-adjoint and positive definite unbounded linear operator with dense domain $\mathcal{D}(B)$, i.e., we have for some constant $m > 0$,

$$\|Bu\| \ge m\|u\| \qquad \text{for } u \in \mathcal{D}(B). \tag{1.5}$$

[†]Faculty of Mathematics, Chemnitz University of Technology, 09107 Chemnitz, Germany (hofmannb@mathematik.tu-chemnitz.de).

[‡]Department of Mathematics, University of Siegen, Walter-Flex-Str. 3, 57068 Siegen, Germany (plato@mathematik.uni-siegen.de).

This allows us to introduce the norms

$$\|u\|_\tau := \|B^\tau u\|, \qquad u \in X_\tau \quad (\tau \in \mathbb{R}),$$

where $X_\tau := \mathcal{D}(B^\tau)$ for $\tau > 0$ and $X_\tau := X$ for $\tau \leq 0$. The fractional powers are defined by means of the resolution of the identity generated by the inverse operator B^{-1}; see, e.g., [7, Section 2.3]. Note that the system of spaces $(X_\tau)_{\tau \in \mathbb{R}}$ equipped with the respective norms is strongly related to the Hilbert scale generated by the operator B; see, e.g., [7, Section 8.4]. However, for $\tau < 0$, a topological completion of the spaces $X_\tau = X$ with respect to the norm $\| \cdot \|_\tau$ is not needed in our setting and thus omitted.

In the present work, we discuss nonlinear Tikhonov regularization (1.3) in particular with an oversmoothing penalty term, where we have $u^\dagger \notin X_1 = \mathcal{D}(B)$, or in other words, $\|u^\dagger\|_1 = +\infty$. This continues some studies started in [11, 12] and [9], where convergence rates and numerical case studies are provided for a priori and a posteriori parameter choices, respectively, under certain smoothness assumptions on $u^\dagger$ and structural conditions on F. Under the same structural conditions, which are also similar to those in the corresponding seminal paper for linear operator equations by Natterer [20], we present—as the novelty of this paper—convergence results based on the Banach–Steinhaus theorem without requiring any smoothness assumptions. The error estimates derived in the context of convergence assertions moreover allow us to prove low-order convergence rates under associated (for example logarithmic) source conditions.

The outline of the remainder is as follows: in Section 2, we introduce Hilbert scales and formulate the basic assumptions, and in addition we establish the well-posedness of Tikhonov regularization used in our setting. Then in Section 3, we introduce auxiliary elements needed for the proof of the convergence results, and in addition we provide preliminary error estimates for Tikhonov regularization, which are based on those auxiliary elements and which are needed for the subsequent convergence proofs. The regularizing properties of an a priori parameter choice as well as a discrepancy principle are considered in Section 4. The suggested discrepancy principle is considered in a form that is suitable for misfit functionals which may depend discontinuously on the regularization parameter $\alpha > 0$. As a byproduct of the derived error estimates, we can prove low-order convergence rates in Section 5. We conclude this paper by presenting some results of numerical experiments.

2. Prerequisites and assumptions.

2.1. Main assumptions. In the following assumption, we briefly summarize the structural properties of the operator F and of its domain $\mathcal{D}(F)$, in particular with respect to the solution $u^\dagger$ of equation (1.3). For examples of nonlinear inverse problems that satisfy these assumptions (or at least substantial parts of it), we refer to [6, 9] and to the appendices of [11, 28].

ASSUMPTION 2.1.
  (a) The operator $F : X \supset \mathcal{D}(F) \to Y$ is sequentially continuous on $\mathcal{D}(F)$ with respect to the weak topologies of the Hilbert spaces X and Y.
  (b) The domain of definition $\mathcal{D}(F) \subset X$ is a closed and convex subset of X.
  (c) Let $\mathcal{D} := \mathcal{D}(F) \cap \mathcal{D}(B) = \mathcal{D}(F) \cap X_1$ be a non-empty set.
  (d) Let the solution $u^\dagger \in \mathcal{D}(F)$ to equation (1.1) with right-hand side $f^\dagger$ be an interior point of the domain $\mathcal{D}(F)$.
  (e) Let the data $f^\delta \in Y$ satisfy the noise model (1.2), and let the initial guess $\overline{u}$ satisfy (1.4).

(f) *Let $a > 0$, and let there exist finite constants $0 < c_a \leq C_a$ such that the inequality chain*

$$(2.1) \qquad c_a\|u - u^\dagger\|_{-a} \leq \|Fu - f^\dagger\| \leq C_a\|u - u^\dagger\|_{-a}$$

holds true for all $u \in \mathcal{D}$.

REMARK 2.2. From (2.1) in item (f) (left-hand side inequality), we have for $u^\dagger \in X_1$ that $u^\dagger$ is the uniquely determined solution to equation (1.1) in the set $\mathcal{D}$. For $u^\dagger \notin X_1$, there is no solution at all to (1.1) in $\mathcal{D}$. But in both cases, alternative solutions $u^* \notin X_1$ with $u^* \in \mathcal{D}(F)$ and $Fu^* = f^\dagger$ cannot be excluded.

2.2. Properties of regularized solutions of Tikhonov regularization. For $\alpha > 0$, minimizers of the Tikhonov functional T_α^δ exist (cf. Proposition 2.4 below) and are denoted by u_α^δ, i.e., we have $T_\alpha^\delta(u_\alpha^\delta) = \min_{u \in \mathcal{D}(F)} T_\alpha^\delta(u)$. Evidently, by definition of the penalty term, $u_\alpha^\delta \in \mathcal{D}$ holds true.

EXAMPLE 2.3. In this example let $F = A : X \to Y$ with $\mathcal{D}(F) = X$ be a bounded linear operator with non-closed range $\mathcal{R}(A)$, and for simplicity let $\overline{u} = 0$. In this setting, Tikhonov regularized solutions u_α^δ solve the linear operator equation

$$(A^*A + \alpha B^{-2})\, u_\alpha^\delta = A^* f^\delta.$$

In the special situation of an injective operator A and of a scale generator $B = (A^*A)^{-q/2}$ with $q > 0$, this gives

$$(A^*A + \alpha(A^*A)^q)\, u_\alpha^\delta = A^* f^\delta,$$

and Assumption 2.1 is then satisfied with $a = 1/q$. The oversmoothing case $u^\dagger \notin X_1$ here means that $u^\dagger \notin \mathcal{R}((A^*A)^{q/2})$. This situation is discussed in the analysis of fractional Tikhonov regularization, and we refer for example to [3, 10, 17]. In Natterer's paper [20], the following analog

$$(2.2) \qquad c_a\|u\|_{-a} \leq \|Au\| \leq C_a\|u\|_{-a} \qquad \text{for all} \quad u \in X$$

of the inequality chain (2.1) is the basis for error estimates and convergence rates results for linear operator equations. The constant $a > 0$ here characterizes the degree of ill-posedness of the problem. For extensions of Natterer's results in the case of linear problems, we refer to [7, Section 8.5]. For nonlinear problems, we mention that Neubauer has discussed in [21] the consequences of the two-sided condition

$$c_a\|u\|_{-a} \leq \|F'(u^\dagger)u\| \leq C_a\|u\|_{-a} \qquad \text{for all} \quad u \in X,$$

which is an extension of (2.2) to the nonlinear case and closely connected to (2.1) if the forward operator F has a Fréchet derivative $F'(u^\dagger)$ at $u^\dagger$. For more details, see also [6, Sect. 4].

The extremal problem (1.3) for finding regularized solutions is well-posed with respect to the existence of minimizers and their stability in a sense specified in the following proposition. This follows by standard results from regularization theory (cf., e.g., [25, Chapter 2.6], [26, 27], and [24, Section 4.1.1]). So we give a sketch of proof only.

PROPOSITION 2.4. *Let Assumption 2.1 be satisfied.*

(a) *There exists, for all $\alpha > 0$, a minimizer u_α^δ of the Tikhonov functional T_α^δ in the set $\mathcal{D}$.*
(b) *Each minimizing sequence of T_α^δ over $\mathcal{D}$ has a subsequence that converges strongly in X_1 to a minimizer $u_\alpha^\delta \in \mathcal{D}$ of the Tikhonov functional.*

(c) *For every $\alpha > 0$, the regularized solutions u_α^δ are stable in X_1 with respect to small perturbations in the data f^δ.*

    Proof. The basic ingredients needed for the proof are as follows:

1. The operator F, when considered as $F : X_1 \supset \mathcal{D} \to Y$, is sequentially continuous with respect to the weak topologies on X_1 and Y. This implies that the misfit functional $u \in \mathcal{D} \mapsto \|Fu - f^\delta\| \in \mathbb{R}$ is sequentially continuous with respect to the weak topology on X_1.

2. The set $\mathcal{D}$ is weakly closed in X_1.

3. The stabilizing functional $\| \cdot - \overline{u}\|_1^2$ is sequentially weakly lower continuous on X_1.

    The statement in the second item follows from the two facts that (i) the embedding operator $X_1 \hookrightarrow X$ is continuous and that (ii) each closed convex subset of a Hilbert space is weakly closed.

    From these ingredients, it follows that each minimizing sequence $(u_n) \subset \mathcal{D}$ of the Tikhonov functional has a subsequence which converges weakly in X_1 to a minimizer u_α^δ, and the corresponding subsequence of $(\|u_n - \overline{u}\|_1)$ converges to $\|u_\alpha^\delta - \overline{u}\|_1$. $\square$

    REMARK 2.5. We note that the minimizer of the Tikhonov functional may be non-unique because T_α^δ can, for nonlinear forward operators F, be a non-convex functional as a consequence of a non-convex misfit term $\|Fu - f^\delta\|^2$. If, for example, $Fu := u \star u$ represents the autoconvolution operator in $X = L^2(0,1)$ (cf., e.g., [5] and references therein) and $\overline{u} = 0$, then we have $T_\alpha^\delta(u) = T_\alpha^\delta(-u)$, which illustrates the non-uniqueness phenomenon. On the other hand, it should be mentioned that the properties of Tikhonov regularization in Hilbert spaces are well investigated when the penalty functional in the Tikhonov functional is replaced by $u \mapsto \|u - \overline{u}\|^2$, cf., e.g., [7, Chapter 10] or [23, Section 3.1] and the references therein, respectively.

    One of the two main goals of this study is to discuss convergence results for Tikhonov regularization with oversmoothing penalty, i.e., $u^\dagger \notin X_1$ (note, however, that this is not explicitly required anywhere), and when the regularization error $u_\alpha^\delta - u^\dagger$ is still measured in the norm of X. This continues former studies like [9] under the assumption $u^\dagger \in X_p$ for some $0 < p < 1$. In contrast to those papers, the focus of the present work is, although also not explicitly required anywhere, on the case $u^\dagger \notin X_p$ for each $0 < p < 1$, thus consequently on the situation characterized by $p = 0$. On the other hand, we also mention convergence assertions for $u^\dagger \in X_p$ with $p \geq 1$ under the inequality chain (2.1).

3. Auxiliary elements and preparatory results.

    3.1. Auxiliary elements. In this section, we consider *auxiliary elements* that are needed to verify our convergence results. As a preparation, we introduce the bounded, injective, selfadjoint, positive semidefinite linear operator

$$(3.1) \qquad\qquad G := B^{-(2a+2)} : X \to X,$$

where the operator B obeying the condition (1.5) is defined in Section 1, and $a > 0$ is introduced in item (f) of Assumption 2.1. Note that the range $\mathcal{R}(G)$ of G is not closed, and hence, zero is an accumulation point of the spectrum $\sigma(G) \subset [0, \|G\|]$ of G. In this context, we also mention that $u \in X_p$ $(p > 0)$ is equivalent to $u \in \mathcal{R}(G^{\frac{p}{2a+2}})$, which means that u obeys a power-type source condition $u = G^{\frac{p}{2a+2}} w$ with some source element $w \in X$. In the case $p = 0$, i.e., if $u \in X$ but $u \notin X_p$ for all $p > 0$, then it was shown in [13, 18] that there exist an index function[1] φ (for example of logarithmic type, cf. [15]) and a source element $w \in X$ such that a (low-order) source condition $u = \varphi(G)w$ is satisfied.

---

[1]According to [19], we call a function $\varphi \colon (0, \infty) \to (0, \infty)$ index function, if it is continuous, non-decreasing, and satisfies the limit condition $\lim_{t \to 0} \varphi(t) = 0$.

The auxiliary elements based on the operator G from (3.1) are defined as follows:

$$(3.2) \quad \widehat{u}_\alpha := \overline{u} + G(G + \alpha I)^{-1}(u^\dagger - \overline{u}) = u^\dagger - \alpha(G + \alpha I)^{-1}(u^\dagger - \overline{u}) \qquad \text{for } \alpha > 0,$$

where the solution $u^\dagger$ of the operator equation (1.1) and the corresponding initial guess $\overline{u}$ are as introduced above. The basic properties of the auxiliary elements are summarized in Lemma 3.1.

We should mention that the auxiliary elements $\widehat{u}_\alpha$ are the uniquely determined minimizers of the *artificial Tikhonov functional*

$$T_{a,\alpha}(u) := \|u - u^\dagger\|_{-a}^2 + \alpha \|u - \overline{u}\|_1^2$$

over all $u \in X$. The mapping $u^\dagger \mapsto \widehat{u}_\alpha$ is a variant of a *proximal operator* and possesses an explicit character. This allows for error estimates and convergence assertions for our Hilbert scale model in the case of oversmoothing penalties. If we leave the Hilbert scales, then it becomes much more difficult to handle oversmoothing penalties. This is exemplified by the work in [8], where the ℓ^1-regularization is studied when the solution $u^\dagger$ is only in ℓ^2. There, the auxiliary elements are constructed by projection mappings instead of proximal mappings, and the occurring conditions for convergence are difficult to interpret.

In order to specify the limit behaviour of different positive functions occurring in the error estimates, in the sequel we use a collection of non-negative functions named $f_i(\alpha)$, $(i = 1, 2, \ldots)$, defined for $\alpha > 0$ and with the property

$$(3.3) \qquad\qquad \lim_{\alpha \to 0} f_i(\alpha) = 0,$$

to be supposed for all indices i. Consequently, we have for all i that $f_i(\alpha) = o(1)$ as $\alpha \to 0$. Note that pairwise products $f_i(\alpha)f_j(\alpha)$ and linear combinations $K_i f_i(\alpha) + K_j f_j(\alpha)$ with non-negative constants K_i, K_j can again be written as such a function $f_k(\alpha) = o(1)$ as $\alpha \to 0$.

LEMMA 3.1. *There are functions $f_i(\alpha)$ ($i = 1, 2, 3, \alpha > 0$) satisfying (3.3) such that the auxiliary elements from (3.2) have the following properties:*
1. $\|\widehat{u}_\alpha - u^\dagger\| = f_1(\alpha) = o(1)$ *as* $\alpha \to 0$,
2. $\|\widehat{u}_\alpha - u^\dagger\|_{-a} = f_2(\alpha)\, \alpha^{a/(2a+2)} = o(\alpha^{a/(2a+2)})$ *as* $\alpha \to 0$,
3. $\|\widehat{u}_\alpha - \overline{u}\|_1 = f_3(\alpha)\, \alpha^{-1/(2a+2)} = o(\alpha^{-1/(2a+2)})$ *as* $\alpha \to 0$.

Proof. We show first that

$$(3.4) \qquad\qquad \|G^\theta(G + \alpha I)^{-1}u\| = o(\alpha^{\theta - 1}) \qquad \text{as } \alpha \to 0$$

holds true for all $0 \le \theta < 1$ and $u \in X$. It is well known that

$$\|G(G + \alpha I)^{-1}\| \le 1, \qquad \|(G + \alpha I)^{-1}\| \le \alpha^{-1} \qquad \text{for } \alpha > 0.$$

Then the interpolation inequality implies the estimate

$$(3.5) \qquad\qquad \|G^\theta(G + \alpha I)^{-1}\| \le \alpha^{\theta - 1} \qquad \text{for } \alpha > 0 \quad \text{and} \quad 0 \le \theta \le 1.$$

Note that the operator G is selfadjoint and positive semidefinite, and thus the fractional powers G^θ are well-defined.

In addition, for fixed $0 \le \theta < 1$ and any $u \in \mathcal{R}(G^q)$ with $q > 0$ chosen so small such that $\theta + q \le 1$, we have, from (3.5) with θ replaced by $\theta + q$,

$$(3.6) \quad \alpha^{1-\theta}\|G^\theta(G + \alpha I)^{-1}u\| = \alpha^{1-\theta}\|G^\theta(G + \alpha I)^{-1}G^q v\| \le \alpha^q \|v\| \to 0 \qquad \text{as } \alpha \to 0,$$

where $u = G^q v$. The asymptotics (3.4) now follows from (3.5) and (3.6) and from an application of the Banach–Steinhaus theorem to the operators $\alpha^{1-\theta} G^\theta (G + \alpha I)^{-1}$ for $\alpha \to 0$. Here $0 \leq \theta < 1$ is fixed, and we have used the fact that for arbitrary $q > 0$, the range of the operator G^q is dense in X, i.e., $\overline{\mathcal{R}(G^q)} = X$. For the Banach–Steinhaus theorem, cf., e.g., [16, Problem 10.1] or [22, Theorem 1.1.4].

For the functions

$$(3.7) \qquad f_1(\alpha) = \|\alpha(G + \alpha I)^{-1}(u^\dagger - \overline{u})\|,$$

$$(3.8) \qquad f_2(\alpha) = \alpha^{-a/(2a+2)}\|G^{a/(2a+2)}[\alpha(G + \alpha I)^{-1}(u^\dagger - \overline{u})]\|,$$

$$(3.9) \qquad f_3(\alpha) = \alpha^{-(2a+1)/(2a+2)}\|G^{(2a+1)/(2a+2)}[\alpha(G + \alpha I)^{-1}(u^\dagger - \overline{u})]\|,$$

the statements of the lemma are now easily obtained from (3.4) and the following three representations,

$$\widehat{u}_\alpha - u^\dagger = -\alpha(G + \alpha I)^{-1}(u^\dagger - \overline{u}),$$
$$B^{-a}(\widehat{u}_\alpha - u^\dagger) = -G^{a/(2a+2)}[\alpha(G + \alpha I)^{-1}(u^\dagger - \overline{u})],$$
$$B(\widehat{u}_\alpha - \overline{u}) = G^{(2a+1)/(2a+2)}[(G + \alpha I)^{-1}(u^\dagger - \overline{u})]. \qquad \square$$

3.2. Some estimates for oversmoothing Tikhonov regularization.

LEMMA 3.2. *Let Assumption 2.1 be satisfied. Then there is a function $f_4(\alpha)$ for $\alpha > 0$ satisfying (3.3) such that for all $\alpha > 0$ and $\delta > 0$, we have*

$$\max\{\|Fu_\alpha^\delta - f^\delta\|, \sqrt{\alpha}\|u_\alpha^\delta - \overline{u}\|_1\} \leq f_4(\alpha)\,\alpha^{a/(2a+2)} + \delta.$$

Proof. For $\alpha > 0$ small enough, say $0 < \alpha \leq \alpha_0$, we have $\widehat{u}_\alpha \in \mathcal{D}$ because item 1 of Lemma 3.1 holds and $u^\dagger$ is an interior point of $\mathcal{D}(F)$. Thus

$$(\|Fu_\alpha^\delta - f^\delta\|^2 + \alpha\|u_\alpha^\delta - \overline{u}\|_1^2)^{1/2} \leq (\|F\widehat{u}_\alpha - f^\delta\|^2 + \alpha\|\widehat{u}_\alpha - \overline{u}\|_1^2)^{1/2}$$
$$\leq \|F\widehat{u}_\alpha - f^\delta\| + \sqrt{\alpha}\|\widehat{u}_\alpha - \overline{u}\|_1 \leq \|F\widehat{u}_\alpha - f^\dagger\| + \sqrt{\alpha}\|\widehat{u}_\alpha - \overline{u}\|_1 + \delta.$$

The first term on the right-hand side of the latter estimate can be written as

$$\|F\widehat{u}_\alpha - f^\dagger\| \leq C_a\|\widehat{u}_\alpha - u^\dagger\|_{-a} \leq C_a\,f_2(\alpha)\,\alpha^{a/(2a+2)}.$$

This is a consequence of item 2 of Lemma 3.1. The second term on the right-hand side of the latter estimate attains the form

$$\sqrt{\alpha}\|\widehat{u}_\alpha - \overline{u}\|_1 \leq f_3(\alpha)\,\alpha^{a/(2a+2)},$$

based on item 3 of Lemma 3.1. This yields the function

$$f_4(\alpha) := C_a\,f_2(\alpha) + f_3(\alpha) \qquad \text{for } \alpha \leq \alpha_0.$$

Note that $f_4(\alpha) \to 0$ as $\alpha \to 0$. For $\alpha > \alpha_0$, the estimate

$$(\|Fu_\alpha^\delta - f^\delta\|^2 + \alpha\|u_\alpha^\delta - \overline{u}\|_1^2)^{1/2} \leq \|F\overline{u} - f^\delta\| \leq \|F\overline{u} - f^\dagger\| + \delta$$

holds, so we may define

$$f_4(\alpha) := \frac{\|F\overline{u} - f^\dagger\|}{\alpha^{a/(2a+2)}} \qquad \text{for } \alpha > \alpha_0.$$

This completes the proof of the lemma. $\qquad \square$

COROLLARY 3.3. *Let Assumption 2.1 be satisfied. Then there are a function $f_5(\alpha)$ for $\alpha > 0$ satisfying (3.3) and a constant $K_1 > 0$ such that for all $\alpha > 0$ and $\delta > 0$, we have*

$$\|u_\alpha^\delta - u^\dagger\|_{-a} \le f_5(\alpha)\,\alpha^{a/(2a+2)} + K_1\,\delta.$$

Proof. It can be concluded from the estimate on the left-hand side of (2.1) and Lemma 3.2 that

$$c_a\|u_\alpha^\delta - u^\dagger\|_{-a} \le \|Fu_\alpha^\delta - f^\dagger\| \le \|Fu_\alpha^\delta - f^\delta\| + \delta \le f_4(\alpha)\,\alpha^{a/(2a+2)} + 2\delta.$$

The assertion of the corollary now follows by setting $f_5(\alpha) := \frac{f_4(\alpha)}{c_a}$ and $K_1 := \frac{2}{c_a}$. □

The error $\|u_\alpha^\delta - u^\dagger\|$ is now bounded by the following series of error estimates. Using the triangle inequality and Lemma 3.1, we obtain

$$(3.10) \qquad \|u_\alpha^\delta - u^\dagger\| \le \|u_\alpha^\delta - \widehat{u}_\alpha\| + \|\widehat{u}_\alpha - u^\dagger\| = \|u_\alpha^\delta - \widehat{u}_\alpha\| + f_1(\alpha),$$

and below we consider the term $\|u_\alpha^\delta - \widehat{u}_\alpha\|$ in more detail. From the interpolation inequality for bounded linear, self-adjoint, and positive semidefinite operators on Hilbert spaces, cf. [7, (2.49)], it follows that

$$(3.11) \qquad \|u_\alpha^\delta - \widehat{u}_\alpha\| \le \|u_\alpha^\delta - \widehat{u}_\alpha\|_{-a}^{1/(a+1)}\|u_\alpha^\delta - \widehat{u}_\alpha\|_1^{a/(a+1)}.$$

Both terms on the right-hand side of the estimate (3.11) can be handled by Corollary 3.3 and Lemma 3.1 in the following manner: Precisely, we find with $f_6(\alpha) := f_2(\alpha) + f_5(\alpha)$ and $f_7(\alpha) := f_3(\alpha) + f_4(\alpha)$ the estimates

$$\|u_\alpha^\delta - \widehat{u}_\alpha\|_{-a} \le \|u_\alpha^\delta - u^\dagger\|_{-a} + \|\widehat{u}_\alpha - u^\dagger\|_{-a} \le f_6(\alpha)\,\alpha^{a/(2a+2)} + K_1\delta,$$

$$\|u_\alpha^\delta - \widehat{u}_\alpha\|_1 \le \|u_\alpha^\delta - \overline{u}\|_1 + \|\widehat{u}_\alpha - \overline{u}\|_1 \le \alpha^{-1/2}\left(f_7(\alpha)\,\alpha^{a/(2a+2)} + \delta\right).$$

Thus, we can continue bounding (3.11). Introducing $f_8(\alpha) := \max\{f_6(\alpha), f_7(\alpha)\}$ and $K_2 := \max\{K_1, 1\}$, we obtain

$$\|u_\alpha^\delta - \widehat{u}_\alpha\| \le \left(f_6(\alpha)\,\alpha^{a/(2a+2)} + K_1\delta\right)^{1/(a+1)}\left(\alpha^{-1/2}\left(f_7(\alpha)\,\alpha^{a/(2a+2)} + \delta\right)\right)^{a/(a+1)}$$

$$\le \left(f_8(\alpha)\,\alpha^{a/(2a+2)} + K_2\delta\right)^{1/(a+1)}\left(\alpha^{-1/2}\left(f_8(\alpha)\,\alpha^{a/(2a+2)} + K_2\delta\right)\right)^{a/(a+1)}$$

$$= \alpha^{-a/(2a+2)}\left(f_8(\alpha)\,\alpha^{a/(2a+2)} + K_2\delta\right) = f_8(\alpha) + K_2\frac{\delta}{\alpha^{a/(2a+2)}}.$$

From the latter estimate and (3.10), the following proposition now immediately follows by considering $f_9(\alpha) := f_1(\alpha) + f_8(\alpha)$ there.

PROPOSITION 3.4. *Let Assumption 2.1 be satisfied. Then there are a function $f_9(\alpha)$ for $\alpha > 0$ satisfying (3.3) and a constant $K_2 > 0$ such that for all $\alpha > 0$ and $\delta > 0$, we have*

$$(3.12) \qquad \|u_\alpha^\delta - u^\dagger\| \le f_9(\alpha) + K_2\frac{\delta}{\alpha^{a/(2a+2)}}.$$

The inequality (3.12), which is valid for arbitrary noise levels $\delta > 0$ and regularization parameters $\alpha > 0$, allows us to formulate in the subsequent section sufficient conditions for the convergence of the error norm $\|u_\alpha^\delta - u^\dagger\|$ of the regularized solutions.

4. Convergence results.

4.1. Main theorem. The following main theorem is an immediate consequence of the error estimates outlined in the preceding section. The formulated convergence result follows straightforwardly from the inequality (3.12).

THEOREM 4.1. *Let Assumption 2.1 be satisfied. Then for any a priori parameter choice $\alpha_* = \alpha(\delta)$ and any a posteriori parameter choice $\alpha_* = \alpha(\delta, y^\delta)$, the regularized solutions $u_{\alpha_*}^\delta$ converge in the norm of the Hilbert space X to the solution $u^\dagger$ of the operator equation (1.1) for $\delta \to 0$, i.e., $\lim_{\delta \to 0} \|u_{\alpha_*}^\delta - u^\dagger\| = 0$ whenever*

$$(4.1) \qquad \alpha_* \to 0 \qquad and \qquad \frac{\delta}{\alpha_*^{a/(2a+2)}} \to 0 \qquad as \qquad \delta \to 0.$$

REMARK 4.2. If[2]

$$(4.2) \qquad \alpha_*^{a/(2a+2)} \sim \delta \qquad as \qquad \delta \to 0,$$

then convergence cannot be derived in this way because in that borderline case, the second term on the right-hand side of inequality (3.12) does not tend to zero.

REMARK 4.3. The convergence result of Theorem 4.1 applies both to (a) the classical case $u^\dagger \in \mathcal{D}(B)$ as well as to (b) oversmoothing penalties $u^\dagger \notin \mathcal{D}(B)$. This theorem is directly based on formula (3.12). An inspection of the proof of this formula, given by means of Lemma 3.2 through Proposition 3.4, shows that both inequalities in (2.1) are needed for the convergence result of Theorem 4.1. Nevertheless, for the oversmoothing case (b), this is real progress since the convergence result of Theorem 4.1 does not use any form of additional smoothness on $u^\dagger$. Such *pure convergence* assertions for the oversmoothing case, without any smoothness assumptions like the condition $u^\dagger \in X_p$ for some $p \in (0, 1)$ that was used in [11] for obtaining *convergence rates* results, are missing in the literature by now. However, as is mentioned in [7, Remark 8.24], general norm convergence of the regularized solutions under the condition (4.1) is known *for linear problems* including the oversmoothing case.

As is well known, in case (a) with $u^\dagger \in \mathcal{D}(B)$, the parameter choice condition

$$\alpha_* \to 0 \qquad and \qquad \frac{\delta^2}{\alpha_*} \to 0 \qquad as \quad \delta \to 0,$$

which is stronger than (4.1), is always sufficient for the convergence of the regularized solutions, and the inequalities occurring in (2.1) represent only tools for obtaining convergence rates. On the other hand, in the limit situation $\alpha_* \sim \delta^2$ of choosing the regularization parameter, one needs the left-hand side inequality in (2.1) for obtaining convergence rates, and this inequality occurs here as a conditional stability estimate (cf. [9, Prop. 3], [6, Theorem 1.1] and the references therein). Convergence is then a consequence of the derived convergence rates.

4.2. A priori parameter choice of power type. In this section, we consider—in light of Theorem 4.1—the a priori parameter choice

$$(4.3) \qquad \alpha_* = \alpha(\delta) \sim \delta^\kappa$$

for some exponent $\kappa > 0$. Then condition (4.1) is satisfied if and only if $0 < \kappa < 2 + \frac{2}{a}$, and the borderline condition (4.2) holds if and only if $\kappa = 2 + \frac{2}{a}$. This yields the following proposition:

---

[2]With a slight abuse of notation, for two nonnegative functions we write $f(\delta) \sim g(\delta)$ if there exist two constants $0 < c_1 \le c_2$ such that $c_1 f(\delta) \le g(\delta) \le c_2 f(\delta)$ for each $\delta > 0$ small enough.

PROPOSITION 4.4. *For the a priori choice* (4.3) *of the regularization parameter* $\alpha > 0$, *condition* (4.1) *in Theorem* 4.1 *holds if and only if* $0 < \kappa < 2 + \frac{2}{a}$. *For all* $a > 0$, *the choice* $\alpha_* \sim \delta^2$ *yields convergence.*

We can distinguish the κ-intervals (A): $0 < \kappa < 2$, (B): $\kappa = 2$, and (C): $2 < \kappa < 2 + \frac{2}{a}$ for (4.3). Then we have $\frac{\delta^2}{\alpha_*} \to 0$ as $\delta \to 0$ in situation (A) and $\frac{\delta^2}{\alpha_*} \sim 1$ in situation (B). Note that both situation also occur and yield convergent regularized solutions in the oversmoothing case $u^\dagger \notin \mathcal{D}(B)$. This is a little bit surprising because the behaviour

$$(4.4) \qquad \frac{\delta^2}{\alpha_*} \to \infty \qquad \text{as} \quad \delta \to 0$$

occurring in situation (C) was supposed in the literature to be typical for the case of over-smoothing penalties. Namely, as is seen in [12], convergence rate results of the form

$$\|u_{\alpha_*}^\delta - u^\dagger\| = \mathcal{O}\left(\delta^{\frac{p}{a+p}}\right) \qquad \text{as} \quad \delta \to 0$$

are obtained under the two-sided structural condition (2.1) and in particular under the smoothness assumption $u^\dagger \in X_p$ for $0 < p < 1$, whenever the a priori parameter choice of type (4.3) with a prescribed exponent $\kappa = \frac{2(a+1)}{a+p} = 2 + \frac{2(1-p)}{a+p}$ applies. Evidently, this prescribed κ satisfies the conditions (4.4) and $2 < \kappa < 2 + \frac{2}{a}$ for all $0 < p < 1$. It is important to note that $p = 0$ coincides with the borderline case $\kappa = 2 + \frac{2}{a}$, which, however, is not sufficient for convergent regularized solutions.

4.3. A discrepancy principle. For the specification of an appropriate discrepancy principle, the behaviour of the misfit functional $\alpha \mapsto \|Fu_\alpha^\delta - f^\delta\|$ needs to be described for $\delta > 0$ fixed. The basic properties are summarized in the following proposition:

PROPOSITION 4.5. *Let Assumption* 2.1 *be satisfied. Then for* $\delta > 0$ *fixed, the function* $\alpha \mapsto \|Fu_\alpha^\delta - f^\delta\|$ *is non-decreasing, with*

$$(4.5) \qquad \lim_{\alpha \to 0} \|Fu_\alpha^\delta - f^\delta\| \le \delta, \qquad \lim_{\alpha \to \infty} \|Fu_\alpha^\delta - f^\delta\| = \|F\overline{u} - f^\delta\|.$$

We have $\lim_{\alpha \to \infty} \|u_\alpha^\delta - \overline{u}\| = 0$.

Proof. We start with the verification of the first statement of the proposition. As a preparation, we show that the function $\alpha \mapsto \|u_\alpha^\delta - \overline{u}\|_1$ is non-increasing. Indeed, for $0 < \alpha \le \beta$ fixed, we have

$$\begin{aligned}
T_\beta^\delta(u_\beta^\delta) &\le T_\beta^\delta(u_\alpha^\delta) = T_\alpha^\delta(u_\alpha^\delta) + (\beta - \alpha)\|u_\alpha^\delta - \overline{u}\|_1^2 \\
&\le T_\alpha^\delta(u_\beta^\delta) + (\beta - \alpha)\|u_\alpha^\delta - \overline{u}\|_1^2 \\
&= T_\beta^\delta(u_\beta^\delta) + (\beta - \alpha)(\|u_\alpha^\delta - \overline{u}\|_1^2 - \|u_\beta^\delta - \overline{u}\|_1^2),
\end{aligned}$$

and thus $\|u_\beta^\delta - \overline{u}\|_1 \le \|u_\alpha^\delta - \overline{u}\|_1$. The first statement of the proposition is now easily obtained: for $0 < \alpha \le \beta$ we have

$$\|Fu_\alpha^\delta - f^\delta\|^2 + \alpha\|u_\alpha^\delta - \overline{u}\|_1^2 = T_\alpha^\delta(u_\alpha^\delta) \le T_\alpha^\delta(u_\beta^\delta) = \|Fu_\beta^\delta - f^\delta\|^2 + \alpha\|u_\beta^\delta - \overline{u}\|_1^2$$
$$\le \|Fu_\beta^\delta - f^\delta\|^2 + \alpha\|u_\alpha^\delta - \overline{u}\|_1^2,$$

and thus $\|Fu_\alpha^\delta - f^\delta\| \le \|Fu_\beta^\delta - f^\delta\|$.

Next we consider the latter statement of the proposition. There holds

$$(4.6) \qquad \|Fu_\alpha^\delta - f^\delta\|^2 + \alpha\|u_\alpha^\delta - \overline{u}\|_1^2 = T_\alpha^\delta(u_\alpha^\delta) \le T_\alpha^\delta(\overline{u}) = \|F\overline{u} - f^\delta\|^2,$$

and thus in particular $\|u_\alpha^\delta - \overline{u}\|_1 = \mathcal{O}(\alpha^{-1/2})$ as $\alpha \to \infty$. The estimate (1.5) implies $\|u_\alpha^\delta - \overline{u}\| = \mathcal{O}(\alpha^{-1/2})$ as $\alpha \to \infty$, which implies the latter statement of the proposition.

The first statement in (4.5) follows directly from Lemma 3.2, and we finally consider the second statement in (4.5). From (4.6) we know that $\lim_{\alpha\to\infty} \|Fu_\alpha^\delta - f^\delta\| \leq \|F\overline{u} - f^\delta\|$. Conversely, sequential weak continuity of the operator F implies the weak convergence $Fu_\alpha^\delta \rightharpoonup F\overline{u}$ as $\alpha \to \infty$, and thus $\lim_{\alpha\to\infty} \|Fu_\alpha^\delta - f^\delta\| \geq \|F\overline{u} - f^\delta\|$. This completes the proof of the proposition. $\quad\square$

REMARK 4.6. Notice that in the proof of Proposition 4.5, no use of the fundamental estimates (2.1) for the smoothing property of F is made. Notice also that the statement of Proposition 4.5 is quite similar to related results for Tikhonov regularization with non-oversmoothing penalty, cf. [1], [25, Section 2.6], and [27, Section 6.7].

It follows from Proposition 4.5 that the following version of the discrepancy principle (cf. [26, 27]) is implementable. It determines, for each noise level $\delta > 0$, an approximation $u_{\alpha_*}^\delta \in \mathcal{D}$. Possible discontinuities of the misfit functional $\alpha \mapsto \|Fu_\alpha^\delta - f^\delta\|$ are taken into account.

ALGORITHM 4.7 (Discrepancy principle). *Let $b > 1$ and $c > 1$ be finite constants.*
  1. *If $\|F\overline{u} - f^\delta\| \leq b\delta$ holds, then choose $\alpha_* = \infty$, i.e., $u_\infty^\delta := \overline{u} \in \mathcal{D}$.*
  2. *Otherwise, choose a finite parameter $\alpha =: \alpha_* > 0$ such that*

$$(4.7) \qquad \|Fu_{\alpha_*}^\delta - f^\delta\| \leq b\delta \leq \|Fu_{\beta_*}^\delta - f^\delta\| \qquad \text{for some } \alpha_* \leq \beta_* \leq c\alpha_*.$$

Algorithm 4.7 can be realized by the following strategy:

REMARK 4.8 (Sequential discrepancy principle). Practically, a parameter α_* satisfying condition (4.7) can be determined, e.g., by choosing a constant $\theta > 1$ and an initial guess $\alpha^{(0)} > 0$ and proceeding then as follows:
  • If $\|Fu_{\alpha^{(0)}}^\delta - f^\delta\| \geq b\delta$ holds, then, with the notation $\alpha^{(k)} = \theta^{-k}\alpha^{(0)}$, proceed for $k = 1, 2, \ldots$ until $\|Fu_{\alpha^{(k)}}^\delta - f^\delta\| \leq b\delta \leq \|Fu_{\alpha^{(k-1)}}^\delta - f^\delta\|$ is satisfied for the first time; define $\alpha_* = \alpha^{(k)}$ then.
  • If $\|Fu_{\alpha^{(0)}}^\delta - f^\delta\| \leq b\delta$ holds, then, with the notation $\alpha^{(k)} = \theta^k \alpha^{(0)}$, proceed for $k = 1, 2, \ldots$ until $\|Fu_{\alpha^{(k-1)}}^\delta - f^\delta\| \leq b\delta \leq \|Fu_{\alpha^{(k)}}^\delta - f^\delta\|$ is satisfied for the first time; define $\alpha_* = \alpha^{(k-1)}$ then.

The regularizing properties of Algorithm 4.7 are stated in the following theorem:

THEOREM 4.9. *Let Assumption 2.1 be satisfied. For the a posteriori parameter choice introduced in Algorithm 4.7, we have*

$$(4.8) \qquad \|u_{\alpha_*}^\delta - u^\dagger\| \to 0, \quad \frac{\delta}{\alpha_*^{a/(2a+2)}} \to 0 \qquad \text{as } \delta \to 0.$$

Proof. For an arbitrary countable noise level set $\Delta \subset \mathbb{R}_+$ having the origin as only accumulation point, we consider the following three cases: (a) $\alpha_* = \infty$ for each $\delta \in \Delta$, (b) $\alpha_* \to 0$ as $\Delta \ni \delta \to 0$, and (c) $\alpha_* < \infty$ for each $\delta \in \Delta$, $\liminf_{\Delta \ni \delta \to 0} \alpha_* > 0$. Below we show that in each of those three cases, (4.8) holds. The main statement of the theorem then follows by a standard subsequence argument. Note that in cases (a) and (c), the second statement in (4.8) trivially holds.

(a) The case $\alpha_* = \infty$ for $\delta \in \Delta$ means $\|F\overline{u} - f^\delta\| \leq b\delta$ for $\delta \in \Delta$, and thus $F\overline{u} = f^\dagger$, and then $u_\infty^\delta = \overline{u} = u^\dagger$ for $\delta \in \Delta$.
(b) Suppose that $\alpha_* \to 0$ as $\Delta \ni \delta \to 0$. From Lemma 3.2, we obtain

$$b\delta \leq \|Fu_{\beta_*}^\delta - f^\delta\| \leq o(\beta_*^{a/(2a+2)}) + \delta = o(\alpha_*^{a/(2a+2)}) + \delta,$$

and thus $\delta/\alpha_*^{a/(2a+2)} \to 0$ as $\Delta \ni \delta \to 0$. Proposition 3.4 then yields $\|u_{\alpha_*}^\delta - u^\dagger\| \to 0$ as $\Delta \ni \delta \to 0$.

(c) Next suppose that both $\alpha_* < \infty$, for $\delta \in \Delta$, and

$$(4.9) \qquad \liminf_{\Delta \ni \delta \to 0} \alpha_* > 0.$$

(i) We first observe that

$$(4.10) \qquad c_a \|u_{\alpha_*}^\delta - u^\dagger\|_{-a} \leq \|Fu_{\alpha_*}^\delta - f^\dagger\| \leq \|Fu_{\alpha_*}^\delta - f^\delta\| + \delta \leq (b+1)\delta,$$

so $\|u_{\alpha_*}^\delta - u^\dagger\|_{-a} = \mathcal{O}(\delta)$ as $\Delta \ni \delta \to 0$. Note that the asymptotics (4.9) is not needed for this result.

(ii) There holds

$$(4.11) \qquad \|u_{\alpha_*}^\delta\|_1 = \mathcal{O}(1) \qquad \text{as } \Delta \ni \delta \to 0.$$

This easily follows from (4.6) and (4.9) in combination with the estimate

$$\sqrt{\alpha_*}\|u_{\alpha_*}^\delta - \overline{u}\|_1 \leq \|F\overline{u} - f^\delta\| \leq \|F\overline{u} - f^\dagger\| + \delta.$$

(iii) We next show that

$$(4.12) \qquad u^\dagger \in \mathcal{D}(B).$$

For this purpose, we observe that the estimate (4.11) implies weak convergence in X_1 for some subsequence $\Delta' \subset \Delta$, i.e., for some element $v \in \mathcal{D}(B) = X_1$, we have $u_{\alpha_*}^\delta \rightharpoonup v$ in X_1 as $\Delta' \ni \delta \to 0$. From the (weak) continuity of the embedding operator $X_1 \hookrightarrow X$, we then obtain $u_{\alpha_*}^\delta \rightharpoonup v$ in X as $\Delta' \ni \delta \to 0$, and thus $v \in \mathcal{D}$ due to the weak closedness of $\mathcal{D}(F)$. Since the operator F is sequentially weakly continuous, we have $Fu_{\alpha_*}^\delta \rightharpoonup Fv$ as $\Delta' \ni \delta \to 0$. Algorithm 4.7 implies $\|Fu_{\alpha_*}^\delta - Fu^\dagger\| \to 0$ as $\delta \to 0$, so we obtain $Fv = Fu^\dagger$. The lower bound in (2.1) then gives $v = u^\dagger$, which finally implies (4.12).

(iv) From (4.10), (4.12), and the interpolation inequality, we now obtain

$$\|u_{\alpha_*}^\delta - u^\dagger\| \leq \|u_{\alpha_*}^\delta - u^\dagger\|_{-a}^{1/(a+1)} \cdot \|u_{\alpha_*}^\delta - u^\dagger\|_1^{a/(a+1)} = \mathcal{O}(\delta^{1/(a+1)}) \qquad \text{as } \Delta \ni \delta \to 0.$$

This completes the proof of the theorem.   □

Note that in the oversmoothing situation $u^\dagger \notin X_1$, case (c) in the proof of Theorem 4.9 does not emerge; cf. (4.12). This fact has, in a similar setting, already been observed in [11, Lemma 1].

REMARK 4.10. Notice that the situation (b) in the proof of Theorem 4.9 is the regular case in applications. The case (c) is an exceptional case, which, in the non-oversmoothing case, can be excluded if the exact penalization veto is satisfied. This veto had been introduced in the paper [1]; see also [2].

5. Low-order convergence rates. Our convergence assertion established in the main theorem formulated in Section 4 is due to the error estimate (3.12) derived in Section 3. The presented sufficient conditions for convergence are based on the Banach–Steinhaus theorem and do not need any form of solution smoothness. In other words, the case $p = 0$ is included, where $u^\dagger$ does not satisfy a power-type source condition. However, as already mentioned above, there exists at least a source condition of lower order for the solution element $u^\dagger \in X$. Precisely, there is always an index function φ and a source element $w \in X$ such that

$$(5.1) \qquad u^\dagger - \overline{u} = \varphi(G)\,w.$$

Based on formula (3.12) and taking into account the representations (3.7), (3.8), and (3.9), we can derive for such a source condition low-order convergence rates in the case of oversmoothing penalties as a byproduct of the studies presented in Section 3. We will outline this in the following.

LEMMA 5.1. *If, for an index function φ, the quotient function $\varphi(t)/t$ is non-increasing for $0 < t \leq \bar{t}$ with some constant $\bar{t} \in (0, \|G\|]$, then there exist positive constants C and $\overline{\alpha}$ such that*

$$(5.2) \qquad \sup_{0 < \lambda \leq \|G\|} \frac{\alpha \varphi(\lambda)}{\lambda + \alpha} \leq C \, \varphi(\alpha) \qquad (0 < \alpha \leq \overline{\alpha}).$$

The assertion of the lemma follows directly from [4, Prop. 3.3].[3]

COROLLARY 5.2. *Let φ be an index function such that for each exponent $\eta > 0$ the quotient function $t^\eta/\varphi(t)$ is strictly increasing for sufficiently small $t > 0$. Then for each $0 \leq \theta < 1$ there exist positive constants C and $\overline{\alpha}$ such that*

$$(5.3) \qquad \sup_{0 < \lambda \leq \|G\|} \frac{\alpha \lambda^\theta \varphi(\lambda)}{\lambda + \alpha} \leq C \, \alpha^\theta \varphi(\alpha) \qquad (0 < \alpha \leq \overline{\alpha}).$$

Proof. We have that for all $0 \leq \theta < 1$, the quotient function $\frac{t^\theta \varphi(t)}{t} = \frac{\varphi(t)}{t^{1-\theta}}$ with $1 - \theta > 0$ is non-increasing for sufficiently small $t > 0$. Consequently, there are, according to formula (5.2) of Lemma 5.1, positive constants C and $\overline{\alpha}$ depending on θ such that (5.3) is valid. ☐

THEOREM 5.3. *Let Assumption 2.1 and the source condition (5.1) be satisfied, where it is supposed that for all $\eta > 0$, the index function φ has a strictly increasing quotient function $t^\eta/\varphi(t)$ for sufficiently small $t > 0$. Then we have, for some positive constant K_0 and K_2 from (3.12) and for all $\delta > 0$ and sufficiently small $\alpha > 0$, the error estimate*

$$(5.4) \qquad \|u_\alpha^\delta - u^\dagger\| \leq K_0 \, \varphi(\alpha) + K_2 \, \frac{\delta}{\alpha^{a/(2a+2)}}.$$

Proof. Based on the source condition (5.1), the functions f_1, f_2, and f_3 from Lemma 3.1 satisfy

$$f_1(\alpha) = \mathcal{O}(\varphi(\alpha)), \quad f_2(\alpha) = \mathcal{O}(\varphi(\alpha)), \quad f_3(\alpha) = \mathcal{O}(\varphi(\alpha)) \qquad \text{as} \quad \alpha \to 0.$$

These properties are immediate consequences of (5.3) taking into account the three representations (3.7), (3.8), and (3.9). Since the function $f_9(\alpha)$ in the error estimate (3.12) can be estimated from above by linear combinations and by maximizing the functions f_1, f_2, and f_3 for $\alpha > 0$ sufficiently small, there is a positive constant K_0 such that $f_9(\alpha) \leq K_0 \, \varphi(\alpha)$ holds for sufficiently small $\alpha > 0$. ☐

This yields directly the following low-order convergence rate result:

COROLLARY 5.4. *Under the assumptions of Theorem 5.3, let $\psi(\alpha) := \varphi(\alpha) \, \alpha^{\frac{a}{2a+2}}$ and $\alpha_* = \alpha(\delta) := \psi^{-1}(\delta)$. Then we have*

$$\|u_{\alpha_*}^\delta - u^\dagger\| = \mathcal{O}\left(\varphi(\psi^{-1}(\delta))\right) \qquad \text{as} \quad \delta \to 0.$$

---

[3]For the understanding of the formula (5.2), the concept of *qualification* for a regularization method introduced in [19] is helpful. Precisely, all index functions $\varphi(t)$ which are covered by the function t are qualifications for the classical Tikhonov regularization applied to the operator G, which implies that an inequality of type (5.2) is valid.

EXAMPLE 5.5. In this example, we consider source conditions (5.1) of logarithmic type with the function

$$(5.5) \qquad \varphi(t) = \varphi_{\log}^{\kappa}(t) := (-\log(t))^{-\kappa} \qquad (\kappa > 0),$$

which is strictly concave for sufficiently small $t > 0$ and can be extended to $(0, \infty)$ as an index function. Is is evident for all $\eta, \kappa > 0$ that the quotient function $t^{\eta}/\varphi_{\log}^{\kappa}(t)$ is strictly increasing for sufficiently small $t > 0$, and Corollary 5.2 applies. This yields the error estimate (5.4) written as

$$\|u_{\alpha}^{\delta} - u^{\dagger}\| \leq K_0 \, (-\log(\alpha))^{-\kappa} + K_2 \, \frac{\delta}{\alpha^{a/(2a+2)}}.$$

For the a priori choice $\alpha_* = \alpha(\delta) \sim \delta^2$ of the regularization parameter, this implies the logarithmic convergence rate

$$(5.6) \qquad \|u_{\alpha_*}^{\delta} - u^{\dagger}\| = \mathcal{O}\left((-\log(\delta))^{-\kappa}\right) = \mathcal{O}\left(\varphi_{\log}^{\kappa}(\delta)\right) \qquad \text{as} \quad \delta \to 0.$$

Note that this parameter choice strategy differs from the one presented in Corollary 5.4.

6. Numerical illustrations. The theoretical results are numerically illustrated for the nonlinear operator $F : \ell_2 \supset \mathcal{D}(F) \to \ell_2$ given by the sum $F = F_1 + F_2$ of a linear operator F_1 and a quadratic operator F_2 as follows:

$$(6.1) \qquad F_1 : \ell_2 \supset \mathcal{D}(F) \to \ell_2, \quad (u_n) \mapsto 7(\tfrac{1}{n}u_n),$$

$$(6.2) \qquad F_2 : \ell_2 \supset \mathcal{D}(F) \to \ell_2, \quad (u_n) \mapsto (\tfrac{1}{n}u_n^2).$$

Here,

$$\mathcal{D}(F) = \{\, u \in \ell_2 \mid \|u\|_{\ell_2} \leq 3 \,\},$$

and $\ell_2 = \{\, (u_n) \mid \|u\|_{\ell_2}^2 = \sum_{n=1}^{\infty} u_n^2 < \infty \,\}$. The stronger norm $\|\cdot\|_1$ is defined by the generator

$$B : \ell_2 \supset \mathcal{D}(B) \to \ell_2, \quad (u_n) \mapsto (nu_n), \qquad \mathcal{D}(B) := \{\, (u_n) \mid (nu_n) \in \ell_2 \,\}.$$

In what follows, we consider the equation $Fu = f^{\dagger}$ having

$$u^{\dagger} = (u_n^{\dagger}), \qquad \text{with} \quad u_1^{\dagger} = 1, \quad u_n^{\dagger} = \frac{1}{\sqrt{n}(\log n)^{2.31}}, \quad n = 2, 3, \ldots,$$

as a solution. Assumption 2.1 is satisfied then; in particular, the two structural inequalities in (2.1) are satisfied for $a = 1$. In addition, we have $u^{\dagger} \notin \mathcal{D}(B)$.

Below, some additional remarks on the numerical tests are given.
- We consider Tikhonov regularization (1.3) with $\overline{u} = 0$.
- For the finite-dimensional approximation needed for the computations, we replace in (6.1), (6.2) the space ℓ_2 by $\mathbb{R}^N$, with $N = 6000$ at each occurrence.
- In the numerical experiments, we consider perturbations of the form $f_n^{\delta} = f_n + \Delta_n$, for $n = 1, 2, \ldots, N$, with uniformly distributed random values Δ_n satisfying the error bound $|\Delta_n| \leq \delta/\sqrt{N}$.

For this framework, we consider Tikhonov regularization (1.3) with an a priori and an a posteriori parameter choice, respectively.

6.1. Numerical results for an a priori parameter choice. We first consider the a priori parameter choice $\alpha_* = \delta^2$ for different values of δ. The numerical results in Table 6.1 confirm the logarithmic convergence rate given by (5.6). Note that a logarithmic-type source condition $u^\dagger = \varphi_{\log}^\kappa(G)w$ is indeed satisfied with $\varphi_{\log}^\kappa(t)$ given by (5.5), which is considered for $t \leq 0.9$ and $\kappa = 1.8$, and

$$w = (w_n), \qquad \text{with} \quad w_n = \frac{4^\kappa}{\sqrt{n}(\log n)^{0.51}}, \quad n = 2, 3, \dots .$$

TABLE 6.1
Numerical results for the a priori parameter choice strategy.

δ	$100 \cdot \delta/\|f\|_{\ell_2}$	$\|u_{\alpha_*}^\delta - u^\dagger\|_{\ell_2}$	$\|u_{\alpha_*}^\delta - u^\dagger\|_{\ell_2} / \varphi_{\log}^\kappa(\delta)$
$8.00 \cdot 10^{-3}$	$7.41 \cdot 10^{-2}$	$5.16 \cdot 10^{-2}$	0.8786
$4.00 \cdot 10^{-3}$	$3.71 \cdot 10^{-2}$	$4.31 \cdot 10^{-2}$	0.9336
$2.00 \cdot 10^{-3}$	$1.85 \cdot 10^{-2}$	$3.65 \cdot 10^{-2}$	0.9776
$1.00 \cdot 10^{-3}$	$9.27 \cdot 10^{-3}$	$3.12 \cdot 10^{-2}$	1.0101
$5.00 \cdot 10^{-4}$	$4.63 \cdot 10^{-3}$	$2.68 \cdot 10^{-2}$	1.0328
$2.50 \cdot 10^{-4}$	$2.32 \cdot 10^{-3}$	$2.32 \cdot 10^{-2}$	1.0457
$1.25 \cdot 10^{-4}$	$1.16 \cdot 10^{-3}$	$2.01 \cdot 10^{-2}$	1.0483
$6.25 \cdot 10^{-5}$	$5.79 \cdot 10^{-4}$	$1.75 \cdot 10^{-2}$	1.0395
$3.12 \cdot 10^{-5}$	$2.90 \cdot 10^{-4}$	$1.51 \cdot 10^{-2}$	1.0193
$1.56 \cdot 10^{-5}$	$1.45 \cdot 10^{-4}$	$1.30 \cdot 10^{-2}$	0.9856
$7.81 \cdot 10^{-6}$	$7.24 \cdot 10^{-5}$	$1.11 \cdot 10^{-2}$	0.9361
$3.91 \cdot 10^{-6}$	$3.62 \cdot 10^{-5}$	$9.26 \cdot 10^{-3}$	0.8671
$1.95 \cdot 10^{-6}$	$1.81 \cdot 10^{-5}$	$7.49 \cdot 10^{-3}$	0.7728

6.2. Numerical results for the discrepancy principle. We next consider the discrepancy principle, cf. Algorithm 4.7, with $b = 4$ and for different values of δ. It is in fact realized by the sequential version considered in Remark 4.8 with $\theta = 10$. The numerical results are reported in Table 6.2. The results in columns 3 and 5 confirm the statement of Theorem 4.9. The results in the last column are presented as an illustration of the asymptotical behavior (4.4).

TABLE 6.2
Numerical results for the discrepancy principle.

δ	$100 \cdot \delta/\|f^\dagger\|_{\ell_2}$	$\|u_{\alpha_*}^\delta - u^\dagger\|_{\ell_2}$	α_*	$\delta/\alpha_*^{1/4}$	δ^2/α_*
$1.00 \cdot 10^{-3}$	$9.27 \cdot 10^{-3}$	$4.09 \cdot 10^{-2}$	$1.00 \cdot 10^{-5}$	$1.78 \cdot 10^{-2}$	0.10
$5.00 \cdot 10^{-4}$	$4.63 \cdot 10^{-3}$	$3.13 \cdot 10^{-2}$	$1.00 \cdot 10^{-6}$	$1.58 \cdot 10^{-2}$	0.25
$2.50 \cdot 10^{-4}$	$2.32 \cdot 10^{-3}$	$2.44 \cdot 10^{-2}$	$1.00 \cdot 10^{-7}$	$1.41 \cdot 10^{-2}$	0.62
$1.25 \cdot 10^{-4}$	$1.16 \cdot 10^{-3}$	$1.92 \cdot 10^{-2}$	$1.00 \cdot 10^{-8}$	$1.25 \cdot 10^{-2}$	1.56
$6.25 \cdot 10^{-5}$	$5.79 \cdot 10^{-4}$	$1.51 \cdot 10^{-2}$	$1.00 \cdot 10^{-9}$	$1.11 \cdot 10^{-2}$	3.91
$3.12 \cdot 10^{-5}$	$2.90 \cdot 10^{-4}$	$1.16 \cdot 10^{-2}$	$1.00 \cdot 10^{-10}$	$9.88 \cdot 10^{-3}$	9.77
$1.56 \cdot 10^{-5}$	$1.45 \cdot 10^{-4}$	$1.17 \cdot 10^{-2}$	$1.00 \cdot 10^{-10}$	$4.94 \cdot 10^{-3}$	2.44
$7.81 \cdot 10^{-6}$	$7.24 \cdot 10^{-5}$	$8.64 \cdot 10^{-3}$	$1.00 \cdot 10^{-11}$	$4.39 \cdot 10^{-3}$	6.10
$3.91 \cdot 10^{-6}$	$3.62 \cdot 10^{-5}$	$5.62 \cdot 10^{-3}$	$1.00 \cdot 10^{-12}$	$3.91 \cdot 10^{-3}$	15.26
$1.95 \cdot 10^{-6}$	$1.81 \cdot 10^{-5}$	$2.48 \cdot 10^{-3}$	$1.00 \cdot 10^{-13}$	$3.47 \cdot 10^{-3}$	38.15

REFERENCES

[1] S. W. ANZENGRUBER, B. HOFMANN, AND P. MATHÉ, *Regularization properties of the sequential discrepancy principle for Tikhonov regularization in Banach spaces*, Appl. Anal., 93 (2014), pp. 1382–1400.

[2] S. W. ANZENGRUBER AND R. RAMLAU, *Morozov's discrepancy principle for Tikhonov-type functionals with nonlinear operators*, Inverse Problems, 26 (2010), Art. 025001, 17 pages.

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

[4] A. BÖTTCHER, B. HOFMANN, U. TAUTENHAHN, AND M. YAMAMOTO, *Convergence rates for Tikhonov regularization from different kinds of smoothness conditions*, Appl. Anal., 85 (2006), pp. 555–578.

[5] S. BÜRGER AND B. HOFMANN, *About a deficit in low-order convergence rates on the example of autoconvolution*, Appl. Anal., 94 (2015), pp. 477–493.

[6] H. EGGER AND B. HOFMANN, *Tikhonov regularization in Hilbert scales under conditional stability assumptions*, Inverse Problems, 34 (2018), Art. 115015, 17 pages.

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

[8] D. GERTH AND B. HOFMANN, *Oversmoothing regularization with ℓ^1-penalty term*, AIMS Math., 4 (2019), pp. 1223–1247.

[9] D. GERTH, B. HOFMANN, AND C. HOFMANN, *Case studies and a pitfall for nonlinear variational regularization under conditional stability*, in Inverse Problems and Related Topics, J. Cheng, S. Lu, and M. Yamamoto, eds., Springer Proceedings in Mathematics & Statistics 310, Springer Nature, Singapore, 2020, pp. 177–203.

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

[11] B. HOFMANN AND P. MATHÉ, *Tikhonov regularization with oversmoothing penalty for non-linear ill-posed problems in Hilbert scales*, Inverse Problems, 34 (2018), Art. 015007, 14 pages.

[12] ———, *A priori parameter choice in Tikhonov regularization with oversmoothing penalty for non-linear ill-posed problems*, in Inverse Problems and Related Topics, J. Cheng, S. Lu, and M. Yamamoto, eds., Springer Proceedings in Mathematics & Statistics 310, Springer Nature, Singapore, 2020, pp. 169–176.

[13] B. HOFMANN, P. MATHÉ, AND H. VON WEIZSÄCKER, *Regularization in Hilbert space under unbounded operators and general source conditions*, Inverse Problems, 25 (2009), Art. 115013, 15 pages.

[14] B. HOFMANN AND R. PLATO, *On ill-posedness concepts, stable solvability and saturation*, J. Inverse Ill-Posed Probl., 26 (2018), pp. 287–297.

[15] T. HOHAGE, *Regularization of exponentially ill-posed problems*, Numer. Funct. Anal. Optim., 21 (2000), pp. 439–464.

[16] R. KRESS, *Linear Integral Equations*, 3rd ed., Springer, New York, 2014.

[17] A. K. LOUIS, *Inverse und schlecht gestellte Probleme*, Teubner, Stuttgart, 1989.

[18] P. MATHÉ AND B. HOFMANN, *How general are general source conditions?*, Inverse Problems, 24 (2008), Art. 015009, 5 pages.

[19] P. MATHÉ AND S. V. PEREVERZEV, *Geometry of linear ill-posed problems in variable Hilbert scales*, Inverse Problems, 19 (2003), pp. 789–803.

[20] F. NATTERER, *Error bounds for Tikhonov regularization in Hilbert scales*, Appl. Anal., 18 (1984), pp. 29–37.

[21] A. NEUBAUER, *Tikhonov regularization of nonlinear ill-posed problems in Hilbert scales*, Appl. Anal., 46 (1992), pp. 59–72.

[22] J. SARANEN AND G. VAINIKKO, *Periodic Integral and Pseudodifferential Equations with Numerical Approximation*, Springer, Berlin, 2002.

[23] O. SCHERZER, M. GRASMAIR, H. GROSSAUER, M. HALTMEIER, AND F. LENZEN, *Variational Methods in Imaging*, Springer, New York, 2009.

[24] T. SCHUSTER, B. KALTENBACHER, B. HOFMANN, AND K. S. KAZIMIERSKI, *Regularization Methods in Banach spaces*, de Gruyter, Berlin, 2012.

[25] A. N. TIKHONOV, A. S. LEONOV, AND A. G. YAGOLA, *Nonlinear Ill-Posed Problems. Vol. 1.*, Chapman & Hall, London, 1998.

[26] G. M. VAINIKKO, *Regularization of ill-posed extremal problems (in Russian)*, in Numerical Methods and Optimization, Valgus Publ., Tallinn, 1988, pp. 56–65.

[27] ———, *Regularisierung nichtkorrekter Aufgaben*, Preprint 200, Fachbereich Mathematik, Universität Kaiserslautern, Kaiserslautern, 1991.

[28] F. WERNER AND B. HOFMANN, *Convergence analysis of (statistical) inverse problems under conditional stability estimates*, Inverse Problems, 36 (2020), Art. 015004, 23 pages.