A PRIORI PARAMETER CHOICE IN TIKHONOV REGULARIZATION WITH OVERSMOOTHING PENALTY FOR NON-LINEAR ILL-POSED PROBLEMS

BERND HOFMANN AND PETER MATHÉ

ABSTRACT. We study Tikhonov regularization for certain classes of non-linear ill-posed operator equations in Hilbert space. Emphasis is on the case where the solution smoothness fails to have a finite penalty value, as in the preceding study *Tikhonov regularization with oversmoothing penalty for non-linear ill-posed problems in Hilbert scales*. Inverse Problems 34(1), 2018, by the same authors. Optimal order convergence rates are established for the specific a priori parameter choice, as used for the corresponding linear equations.

1. INTRODUCTION

The present paper is closely related to the recent work [4] of the authors published in the journal *Inverse Problems* devoted to the Tikhonov regularization for non-linear operator equation with oversmoothing penalties in Hilbert scales. Here we adopt the model and terminology. Since ibidem convergence rates of optimal order were only proven for the discrepancy principle as an a posteriori choice of the regularization parameter, we try here to close a gap in regularization theory by extending the same rate results to the case of appropriate a priori choices. This is in good coincidence with the corresponding results for linear operator equations presented in the seminal paper [8], where the same a priori parameter choice successfully allowed for order optimal convergence rates also in the case of oversmoothing penalties.

We consider the approximate solution of an operator equation

$$F(x) = y, \tag{1}$$

which models an inverse problem with an (in general) non-linear forward operator $F : \mathcal{D}(F) \subseteq X \to Y$ mapping between the infinite dimensional Hilbert spaces X and Y, with domain $\mathcal{D}(F)$. By $x^\dagger$ we denote a solution to (1) for given right-hand side y. As a consequence of the 'smoothing' property of F, which is typical for inverse problems, the non-linear equation (1) is *locally ill-posed* at the solution point $x^\dagger \in \mathcal{D}(F)$ (cf. [5, Def. 2]), which in particular means that stability estimates of the form

$$\|x - x^\dagger\|_X \le \varphi(\|F(x) - F(x^\dagger)\|_Y) \tag{2}$$

cannot hold for all $x \in \mathcal{D}(F)$ in an arbitrarily small ball around $x^\dagger$ and for strictly increasing continuous functions $\varphi : [0, \infty) \to [0, \infty)$ with $\varphi(0) = 0$. However, inequalities similar to (2), called *conditional stability estimates*, can hold on the one

---

hand if the admissible range of $x \in \mathcal{D}(F)$ is restricted to densely defined subspaces of X. In this context, we refer to the seminal paper [2] as well as to [1, 6, 7] and references therein. On the other hand, they can hold for all $x \in \mathcal{D}(F)$ if the term $\|x - x^\dagger\|_X$ on the left-hand side of the inequality (2) is replaced with a weaker distance measure, for example a weaker norm (cf. [3] and references therein). In this paper, we follow a combination of both approaches in a Hilbert scale setting.

Based on noisy data $y^\delta \in Y$, obeying the deterministic noise model

$$(3) \qquad\qquad \|y - y^\delta\|_Y \le \delta$$

with *noise level* $\delta > 0$, we use within the domain

$$\mathcal{D} := \mathcal{D}(F) \cap \mathcal{D}(B) \neq \emptyset$$

minimizers $x_\alpha^\delta \in \mathcal{D}$ of the Tikhonov functional T_α^δ solving the extremal problem

$$(4) \qquad T_\alpha^\delta(x) := \|F(x) - y^\delta\|_Y^2 + \alpha\|B(x - \bar{x})\|_X^2 \to \min, \ \text{ subject to } \ x \in \mathcal{D},$$

as stable approximate solutions (regularized solutions) to $x^\dagger$. Above, the element $\bar{x} \in \mathcal{D}$ is a given smooth reference element, and $B \colon \mathcal{D}(B) \subset X \to X$ is a densely defined, unbounded, linear self-adjoint operator, which is strictly positive such that we have for some $m > 0$

$$(5) \qquad\qquad \|Bx\|_X \ge m\|x\|_X, \quad \text{for all} \quad x \in \mathcal{D}(B).$$

This operator B generates a *Hilbert scale* $\{X_\tau\}_{\tau \in \mathbb{R}}$ with $X_0 = X$, $X_\tau = \mathcal{D}(B^\tau)$, and with corresponding norms $\|x\|_\tau := \|B^\tau x\|_X$.

Here we shall focus on the case of an oversmoothing penalty, which means that $x^\dagger \notin \mathcal{D}(B)$ such that $T_\alpha^\delta(x^\dagger) = \infty$. In this case, the regularizing property $T_\alpha^\delta(x_\alpha^\delta) \le T_\alpha^\delta(x^\dagger)$ does not provide additional value here. Throughout this paper, we assume the operator F to be sequentially weakly continuous and its domain $\mathcal{D}(F)$ to be a convex and closed subset of X, which makes the general results of [9, Section 4.1.1] on existence and stability of the regularized solutions x_α^δ applicable.

The paper is organized as follows: In Section 2 we formulate non-linearity and smoothness assumptions, which are required for obtaining a convergence rate result for Tikhonov regularization in Hilbert scales in the case of oversmoothing penalties under an appropriate a priori choice of the regularization parameter. Also in Section 2 we formulate the main theorem. Its proof will then follow from two propositions which are stated. Section 3 is devoted to proving both propositions. Section 4 completes the paper with some concluding discussions.

2. Assumptions and main result

In accordance with the previous study [4] we make the following additional assumption on the structure of non-linearity for the forward operator F with respect to the Hilbert scale generated by the operator B. Sufficient conditions and examples for this non-linearity assumption can be found in the appendix of [4].

Assumption 1 (Non-linearity structure)**.** *There is a number $a > 0$, and there are constants $0 < c_a \le C_a < \infty$ such that*

$$(6) \qquad c_a\|x - x^\dagger\|_{-a} \le \|F(x) - F(x^\dagger)\|_Y \le C_a\|x - x^\dagger\|_{-a} \quad \text{for all } x \in \mathcal{D}.$$

The left-hand inequality of condition (6) implies that, for the right-hand side y, there is no solution to (1) which belongs to $\mathcal{D}$. Moreover note that the parameter

$a > 0$ in Assumption 1 can be interpreted as *degree of ill-posedness* of the mapping F at $x^\dagger$.

The solution smoothness is measured with respect to the generator B of the Hilbert scale as follows. We fix the reference element $\bar{x} \in \mathcal{D}$, occurring in the penalty functional of T_α^δ.

Assumption 2 (Solution smoothness). *There are $0 < p < 1$ and $E < \infty$ such that $x^\dagger \in \mathcal{D}(B^p)$ and*

$$(7) \qquad x^\dagger - \bar{x} \in M_{p,E} := \{x \in X_p, \quad \|x\|_p := \|B^p x\|_X \leq E\}.$$

Moreover, we assume that $x^\dagger$ is an interior point of $\mathcal{D}(F)$, but $x^\dagger \notin \mathcal{D}(B)$.

We shall analyze the error behavior of the minimizer x_α^δ to the Tikhonov functional T_α^δ for the following specific *a priori* parameter choice.

Assumption 3 (A priori parameter choice). *Given noise level $\delta > 0$, ill-posedness degree $a > 0$ (cf. Assumption 1) and solution smoothness $p \in (0,1)$ (cf. Assumption 2), let*

$$(8) \qquad \alpha_* = \alpha_*(\delta) := \delta^{\frac{2(a+1)}{a+p}}.$$

We shall occasionally use the identity $\frac{\delta}{\sqrt{\alpha_*}} = \delta^{\frac{p-1}{a+p}}$, and we highlight that for this parameter choice we have $\frac{\delta}{\sqrt{\alpha_*(\delta)}} \to \infty$ as $\delta \to 0$, for $0 < p < 1$.

The main result is as follows.

Theorem 1. *Under the assumptions stated above let $x_{\alpha_*}^\delta$ be the minimizer of the Tikhonov functional $T_{\alpha_*}^\delta$ for the a priori choice α_* from (8). Then we have the convergence rate*

$$(9) \qquad \|x_{\alpha_*}^\delta - x^\dagger\|_X = \mathcal{O}\left(\delta^{\frac{p}{a+p}}\right) \quad as \quad \delta \to 0.$$

This asymptotics is an immediate consequence of the following two propositions, the proofs of which will be given in the next section.

Proposition 1. *Under the a priori choice α_* from (8) and for sufficiently small $\delta > 0$ we have that*

$$(10) \qquad \|x_{\alpha_*}^\delta - x^\dagger\|_{-a} \leq K\delta$$

holds for some positive constant K.

Proposition 2. *Under the a priori choice α_* from (8) and for sufficiently small $\delta > 0$ we have that*

$$(11) \qquad \|x_{\alpha_*}^\delta - x^\dagger\|_p \leq \tilde{E}$$

holds for some positive constant $\tilde{E}$.

Proof of Theorem 1. Taking into account the assertions of Propositions 1 and 2, the convergence rate (9) follows directly from the interpolation inequality of the Hilbert scale $\{X_\tau\}_{\tau \in \mathbb{R}}$, applied here in the form

$$(12) \qquad \|x_{\alpha_*}^\delta - x^\dagger\|_X \leq \|x_{\alpha_*}^\delta - x^\dagger\|_{-a}^{\frac{p}{a+p}} \|x_{\alpha_*}^\delta - x^\dagger\|_p^{\frac{a}{a+p}}.$$

Thus, the proof of the theorem is complete. $\qquad \square$

3. Proofs of Propositions 1 and 2

Propositions 1 and 2 yield bounds in the (weak) $(-a)$-norm and in the (strong) (p)-norm, respectively. For the proofs we shall use *auxiliary elements* x_{α_*}, constructed as follows. Precisely, in conjunction with the Tikhonov functional T_α^δ from (4) we consider the artificial Tikhonov functional

$$(13) \qquad T_{-a,\alpha}(x) := \|x - x^\dagger\|_{-a}^2 + \alpha\|B(x - \bar{x})\|_X^2,$$

which is well-defined for all $x \in X$. Let x_α be the minimizers of $T_{-a,\alpha}$ over all $x \in X$ that are for all $\alpha > 0$ independent of the noise level $\delta > 0$ and recall now the parameter choice (8) from Assumption 3. For this choice of the regularization parameter the estimates from [4, Prop. 2] yield immediately the following assertions.

Proposition 3. *Suppose that* $x^\dagger \in M_{p,E}$ *for some* $0 < p < 1$, *and let* x_α *be the minimizer of* $T_{-a,\alpha}$. *Given* $\alpha_* > 0$ *as in Assumption 3 the resulting element* x_{α_*} *obeys the bounds*

$$(14) \qquad \|x_{\alpha_*} - x^\dagger\|_X \le E\delta^{p/(a+p)},$$

$$(15) \qquad \|B^{-a}(x_{\alpha_*} - x^\dagger)\|_X \le E\delta,$$

$$(16) \qquad \|B(x_{\alpha_*} - \bar{x})\|_X \le E\delta^{(p-1)/(a+p)} = E\frac{\delta}{\sqrt{\alpha_*}},$$

and

$$(17) \qquad T_{-a,\alpha}(x_{\alpha_*}) \le 2E^2\delta.$$

Moreover, we have that

$$(18) \qquad \|x_{\alpha_*} - \bar{x}\|_p \le E, \ \text{ and } \|x_{\alpha_*} - x^\dagger\|_p \le E.$$

Notice, that in contrast to the solution element $x^\dagger$ the auxiliary element x_{α_*} belongs to $\mathcal{D}$, provided that δ is small enough, and hence we can use the minimizing property

$$(19) \qquad T_{\alpha_*}^\delta(x_{\alpha_*}^\delta) \le T_{\alpha_*}^\delta(x_{\alpha_*}).$$

We derive the following consequence of Proposition 3 formulated in Proposition 4.

Proposition 4. *Let* $x_{\alpha_*}^\delta$ *be the minimizer of* $T_{\alpha_*}^\delta$ *for the Tikhonov functional* T_α^δ *from (4) with the choice* α_*, *as in Assumption 3, of the regularization parameter* $\alpha > 0$. *Then we have for sufficiently small* $\delta > 0$ *that*

$$(20) \qquad \|F(x_{\alpha_*}^\delta) - y^\delta\|_Y \le C\delta,$$

and

$$(21) \qquad \|B(x_{\alpha_*}^\delta - \bar{x})\|_X \le C\frac{\delta}{\sqrt{\alpha_*}},$$

where $C := \left((C_a E + 1)^2 + E^2\right)^{1/2}$.

Proof. Using (19) it is enough to bound $T_{\alpha_*}^\delta(x_{\alpha_*})$ by $C\delta$. We first argue that x_{α_*}, the minimizer of the auxiliary functional (13) for $\alpha = \alpha_*$, belongs to the set $\mathcal{D}(F) \cap \mathcal{D}(B)$. Indeed, the bound (16) shows that $x_{\alpha_*} \in \mathcal{D}(B)$, and the bound (14) indicates that $x_{\alpha_*} \in \mathcal{D}(F)$ for sufficiently small $\delta > 0$, where we use that $x^\dagger$ is an

interior point of $\mathcal{D}(F)$. Now we find from (15) & (16), and from the right-hand inequality of (6), that

$$
\begin{aligned}
T_{\alpha_*}^\delta(x_{\alpha_*}) &\leq \left(\|F(x_{\alpha_*}) - F(x^\dagger)\|_Y + \|F(x^\dagger) - y^\delta\|_Y \right)^2 + \alpha_* \|B(x_{\alpha_*} - \bar{x})\|_X^2 \\
&\leq \left(C_a \|x_{\alpha_*} - x^\dagger\|_{-a} + \delta \right)^2 + E^2 \alpha_* \delta^{2(p-1)/(a+p)} \\
&\leq (C_a E \delta + \delta)^2 + E^2 \delta^2 \\
&= \left((C_a E + 1)^2 + E^2 \right) \delta^2.
\end{aligned}
$$

This completes the proof of Proposition 4. □ □

Now we turn to the proofs of Propositions 1 and 2.

Proof of Proposition 1. Here we use the left-hand side in the non-linearity condition from Assumption 1 and find for sufficiently small $\delta > 0$

$$
\begin{aligned}
\|x_{\alpha_*}^\delta - x^\dagger\|_{-a} &\leq \frac{1}{c_a} \|F(x_{\alpha_*}^\delta) - F(x^\dagger)\|_Y \\
&\leq \frac{1}{c_a} \left(\|F(x_{\alpha_*}^\delta) - y^\delta\|_Y + \|F(x^\dagger) - y^\delta\|_Y \right) \\
&\leq \frac{1}{c_a} (C\delta + \delta) = \frac{1}{c_a} (C + 1)\delta = K\delta,
\end{aligned}
$$

where C is the constant from Proposition 4 and $K := \frac{1}{c_a}(C+1)$. The proof is complete. □

In order to establish the bound occurring in Proposition 2 we start with the following estimate.

Lemma 1. *Let α_* be as in Assumption 3. Then there is a constant $\tilde{C}$ such that we have for sufficiently small $\delta > 0$ that*

$$
\|B(x_{\alpha_*}^\delta - x_{\alpha_*})\|_X \leq \tilde{C} \frac{\delta}{\sqrt{\alpha_*}}.
$$

Proof. By using the triangle inequality we find that

$$
\|B(x_{\alpha_*}^\delta - x_{\alpha_*})\|_X \leq \|B(x_{\alpha_*} - \bar{x})\|_X + \|B(x_{\alpha_*}^\delta - \bar{x})\|_X
$$

The first summand on the right was bounded in (16), and the second was bounded in Proposition 4 for sufficiently small $\delta > 0$. This yields the assertion with $\tilde{C} := C + E$, where C is the constant from Proposition 4. □ □

This allows for the following result.

Proposition 5. *There is a constant $\bar{E}$ such that we have for α_* as in Assumption 3 and for sufficiently small $\delta > 0$ that*

$$
\|x_{\alpha_*}^\delta - x_{\alpha_*}\|_p \leq \bar{E}.
$$

Proof. Again, we use the interpolation inequality, now in the form of

$$
\|x_{\alpha_*}^\delta - x_{\alpha_*}\|_p \leq \|x_{\alpha_*}^\delta - x_{\alpha_*}\|_1^{\frac{a+p}{a+1}} \|x_{\alpha_*}^\delta - x_{\alpha_*}\|_{-a}^{\frac{1-p}{a+1}}.
$$

The norm in the first factor was bounded in Lemma 1. The norm in the second factor can be bounded from

$$
(22) \qquad \|x_{\alpha_*}^\delta - x_{\alpha_*}\|_{-a} \leq \|x_{\alpha_*}^\delta - x^\dagger\|_{-a} + \|x_{\alpha_*} - x^\dagger\|_{-a}.
$$

Now we discuss both summands in the right-hand side of the inequality (22). The first summand was bounded in Proposition 1 by a multiple of δ, and the same holds true for the second summand by virtue of (15). Therefore, there is a constant $\bar{E}$ such that

$$\|x_{\alpha_*}^\delta - x_{\alpha_*}\|_p \le \bar{E}\left(\frac{\delta}{\sqrt{\alpha_*}}\right)^{\frac{a+p}{a+1}} \delta^{\frac{1-p}{a+1}} = \bar{E}\,\delta\,\alpha_*^{-\frac{a+p}{2(a+1)}} = \bar{E}.$$

This completes the proof of Proposition 5.   □

We have gathered all auxiliary estimates in order to turn to the final proof.

Proof of Proposition 2. The estimate (11) is an immediate consequence of Proposition 5 and of the second bound from (18), overall yielding the constant $\tilde{E} := \bar{E} + E$. This completes the proof.   □

4. Conclusions

We have shown that under the non-linearity Assumption 1 and the solution smoothness given as in Assumption 2 the a priori regularization parameter choice

$$\alpha_* = \alpha_*(\delta) = \delta^{\frac{2(a+1)}{a+p}} = \delta^{2-\frac{2(p-1)}{a+p}}$$

allows for the order optimal convergence rate (9) for all $0 < p < 1$. The obtained rate from Theorem 1 is valid for $0 < p \le a+1$ when using the same a priori choice of the regularization parameter from Assumption 3. In all cases, we have that $\alpha_*(\delta) \to 0$ as $\delta \to 0$. In the regular case with $1 < p \le a+1$ we also find the usual convergence $\frac{\delta^2}{\alpha(\delta)} \to 0$ as $\delta \to 0$. For the borderline case $p = 1$ the quotient $\frac{\delta^2}{\alpha_*(\delta)}$ is constant. However, in the oversmoothing case $0 < p < 1$, as considered here, we find that $\frac{\delta^2}{\alpha_*(\delta)} \to \infty$ as $\delta \to 0$.

We stress another observation. In the regular case with $p > 1$ we have the *convergence property*

$$(23) \qquad\qquad \lim_{\delta \to 0} \|x_{\alpha_*}^\delta - x^\dagger\|_X = 0\,,$$

which is a consequence of the sequential weak compactness of a ball and of the Kadec-Klee property in the Hilbert space X. This cannot be shown for $0 < p < 1$ without using additional smoothness of the solution $x^\dagger$. Hence, the convergence property (23) as an implication of Theorem 1 is essentially based on the existence of a positive value p in (7) expressing solution smoothness.

Acknowledgment

References

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

[2] J. Cheng and M. Yamamoto, *One new strategy for a priori choice of regularizing parameters in Tikhonov's regularization*, Inverse Problems **16** (2000), no. 4, L31–L38.

[3] H. Egger and B. Hofmann, *Tikhonov regularization in Hilbert scales under conditional stability assumptions*, Inverse Problems **34** (2018), no. 11, 115015 (17pp).

[4] B. Hofmann and P. Mathé, *Tikhonov regularization with oversmoothing penalty for non-linear ill-posed problems in Hilbert scales*, Inverse Problems **34** (2018), no. 1, 015007 (14pp).

[5] B. Hofmann and O. Scherzer, *Factors influencing the ill-posedness of nonlinear problems*, Inverse Problems **10** (1994), no. 6, 1277–1297.

[6] B. Hofmann and M. Yamamoto, *On the interplay of source conditions and variational inequalities for nonlinear ill-posed problems*, Appl. Anal. **89** (2010), no. 11, 1705–1727.

[7] T. Hohage and F. Weidling, *Variational source conditions and stability estimates for inverse electromagnetic medium scattering problems*, Inverse Probl. Imaging **11** (2017), no. 1, 203–220.

[8] F. Natterer, *Error bounds for Tikhonov regularization in Hilbert scales*, Applicable Anal. **18** (1984), no. 1-2, 29–37.

[9] T. Schuster, B. Kaltenbacher, B. Hofmann, and K. S. Kazimierski, *Regularization Methods in Banach Spaces*, Radon Series on Computational and Applied Mathematics, vol. 10, Walter de Gruyter, Berlin/Boston, 2012.

Faculty of Mathematics, Chemnitz University of Technology, 09107 Chemnitz, Germany
E-mail address: bernd.hofmann@tu-chemnitz.de

Weierstrass Institute for Applied Analysis and Stochastics, Mohrenstrasse 39, 10117 Berlin, Germany
E-mail address: peter.mathe@wias-berlin.de