Weierstraß-Institut
für Angewandte Analysis und Stochastik

im Forschungsverbund Berlin e.V.

Preprint ISSN 0946 – 8633

Interpolation in variable Hilbert scales with application to inverse problems

Peter Mathé and Ulrich Tautenhahn

submitted: 29. June 2006

No. 1148
Berlin 2006

2000 *Mathematics Subject Classification.* 65J20 secondary: 46B70, 65R30.

Key words and phrases. regularization, interpolation spaces, Hilbert scales.

Edited by
Weierstraß-Institut für Angewandte Analysis und Stochastik (WIAS)
Mohrenstraße 39
10117 Berlin
Germany

Fax: + 49 30 2044975
E-Mail: preprint@wias-berlin.de
World Wide Web: http://www.wias-berlin.de/

ABSTRACT. For solving linear ill-posed problems with noisy data regularization methods are required. In the present paper regularized approximations in Hilbert scales are obtained by a general regularization scheme. The analysis of such schemes is based on new results for interpolation in Hilbert scales. Error bounds are obtained under general smoothness conditions.

1. INTRODUCTION

Ill-posed problems arise in several contexts and have important applications in science and engineering (see, e.g., [7, 11, 13, 30]). In this paper we consider ill-posed problems

$$Ax = y \tag{1.1}$$

with bounded linear operators $A : X \to Y$ mapping between infinite dimensional Hilbert spaces X and Y with norms $\| \cdot \|$. Throughout this paper we assume that A is injective and that the range $\mathcal{R}(A)$ is not closed, such that (1.1) has a unique solution $x^\dagger \in X$. If instead of the exact right hand side y we have only noisy data $y^\delta \in Y$ with

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

and known noise level δ, then regularization methods are required in order to obtain regularized approximations that depend continuously on the data. By Tikhonov regularization in Hilbert scales, the regularized approximation x_α^δ is obtained as the solution of the minimization problem $J_\alpha(x) \to MIN$, of the function

$$J_\alpha(x) = \|Ax - y^\delta\|^2 + \alpha\|B^s x\|^2, \qquad x \in \mathcal{D}(B^s), \tag{1.3}$$

where the mapping $B : \mathcal{D}(B) \subseteq X \to X$ is some unbounded densely defined self-adjoint strictly positive definite operator, s is in general some non-negative real number and $\alpha > 0$ is some regularization parameter to be chosen properly.

In many practical problems the constraint operator B which influences the properties of the regularized approximation is chosen to be a differential operator in some appropriate function space, e.g., L^2-space. Natterer has shown in [25] that under the assumptions

$$\|B^p x^\dagger\| \le E \quad \text{and} \quad m\|B^{-a}x\| \le \|Ax\| \le M\|B^{-a}x\| \tag{1.4}$$

with some constants E, m and M, the Tikhonov regularized approximation x_α^δ of problem (1.3) guarantees order optimal error bounds

$$\|x_\alpha^\delta - x^\dagger\| = \mathcal{O}(\delta^{p/(a+p)}) \quad \text{for} \quad s \ge (p-a)/2, \tag{1.5}$$

in case α is chosen *a priori* by $\alpha \asymp \delta^{2(a+s)/(a+p)}$.

Since then regularization in Hilbert scales became attractive, see e.g., [22, 26], where the method (1.3) has been studied with α chosen according to Morozov's discrepancy principle, [23, 28] where this has been generalized to a general regularization scheme, and [23, 25], where extensions to the case of infinitely smoothing operators A have been treated or finally [7, 12, 27, 29], in which extensions to the nonlinear case may be found. We emphasize that interpolation inequalities are used to establish such results and we refer to [7, Chapt. 8.4] for details.

The aim of this paper is to derive order optimal convergence rate results for a general class of regularization methods under smoothness assumptions that are more general than (1.4), thus extending the analysis from [7, Chapt. 8.4]. We first introduce the setup and establish a lower bound as a benchmark. The basic ingredient is some new result on interpolation of operators, which is established in § 3. The present analysis accompanies the one from [3], where the special case $s = 0$ has been treated, and the one from [20], as we regard the problem under the different perspective of operator theory, in particular using factorization and interpolation of operators. Section 4 deals with preliminary considerations, providing the tools which will enable us to obtain error bounds in various norms in § 5. These results are finally used to obtain order optimal error bounds for the regularization schemes under consideration. In § 6 we study *a posteriori* rules for choosing the regularization parameter. We conclude this study in § 7 with a brief discussion of the relevance of the underlying asumptions.

2. SETUP, LOWER BOUND

Throughout we shall measure the smoothness of the unknown solution $x^\dagger$ in terms of some densely defined unbounded self-adjoint strictly positive operator $B\colon \mathcal{D}(B) \subset X \to X$. For technical reasons we shall work with its inverse $G := B^{-1}$ which is a bounded linear injective and self-adjoint operator with non-closed range $\mathcal{R}(G)$. Given constants $E < \infty$ and $p > 0$ we suppose throughout that $x^\dagger \in G_p(E)$ with

$$G_{p,E} = \{x = G^p v, \quad v \in X \text{ with } \|v\| \le E\}\,. \tag{2.1}$$

Remark 1. We mention that we could also require smoothness given in terms of general smoothness conditions, as e.g., $x^\dagger \in G_\tau(E)$, where the proper index function τ, introduced below, replaces $t \mapsto t^p$ in (2.1). Under similar assumptions on operator monotonicity the reader easily extends our results to this more general setup, just by replacing powers of p by application of the function τ.

In the analysis the interplay between the operator G in which solution smoothness is measured and the operator A governing the equation is crucial. As in [18] we call a (continuous) function $\phi : [0, b] \to \mathbb{R}^+$ for some $b > 0$ an *index function*. It is called *proper*, if it is increasing and obeys $\phi(0+) = 0$. Consequently we have $\phi(t) > 0$ provided $t > 0$.

The following set of assumptions extends the *link assumption* in (1.4) to the present setup. Let ϱ be a proper index function.

Assumption A.1. There exists some constant $m > 0$ such that

$$m\|\varrho(G)x\| \le \|Ax\|, \quad x \in X. \tag{2.2}$$

Assumption A.2. There exists some constant $M > 0$ such that

$$\|Ax\| \le M\|\varrho(G)x\|, \quad x \in X. \tag{2.3}$$

Note that necessarily $m \le M$.

Before turning to the problem of regularized approximation to the ill-posed equation (1.1) based on data y^δ as in (1.2), we state the following lower bound, which serves as a benchmark for the accuracy, which can be achieved by any kind of regularization. Let $R : Y \to X$ be an arbitrary mapping and $R(y^\delta)$ be an approximate solution for $x^\dagger$ based on y^δ. The quantity

$$\Delta(G_{p,E}, R, \delta) = \sup \left\{ \|R(y^\delta) - x^\dagger\|,\ y^\delta \in Y,\ \|y - y^\delta\| \le \delta,\ x^\dagger \in G_{p,E} \right\}$$

is called the *worst case error* for identifying $x^\dagger \in G_{p,E}$. This is related to the *modulus of continuity*

$$\omega(G_{p,E}, \delta) := \sup \left\{ \|x\|, \quad x \in G_{p,E},\ \|Ax\| \le \delta \right\},\ \delta > 0. \tag{2.4}$$

It is well known that $\Delta(G_{p,E}, R, \delta) \ge \omega(G_{p,E}, \delta)$ for any $R\colon Y \to X$, and we mention the following result from [24, proof of Thm. 2.2]. For its formulation the functions, defined on $(0, b]$ for some $b \ge \|G\|$ as

$$\Psi_s(t) := t^s \varrho(t), \quad 0 < t \le b, \quad s \in \mathbb{R}, \tag{2.5}$$

turn out to be important.

Theorem 2.1. *If δ/ME belongs to the spectrum of the operator $\Psi_p(G)$ and if Assumption A.2 holds true then*

$$\inf_R \Delta(G_{p,E}, R, \delta) \ge \omega(G_{p,E}, \delta) \ge E \left[\Psi_p^{-1} \left(\frac{\delta}{ME} \right) \right]^p . \tag{2.6}$$

3. Interpolation in variable Hilbert scales

Since the initial study [25] of inverse problems in Hilbert scales it became evident that *interpolation* properties are a main source for obtaining tight error bounds. This is easily seen from the definition of the modulus of continuity in (2.4), which can be understood as relating the norm $\|x\|$ in the original space X to the stronger norm $\|x\|_p := \|G^{-p}x\|$ and the weaker $\|(A^*A)^{1/2}x\|$. This type of interpolation inequalities was carried over to spaces given through general source conditions and we recall the definition of such spaces. If G is a non-negative operator in X and φ is any proper index function, then the set $B_\varphi = \{x = \varphi(G)v, \quad v \in X,\ \|v\| \le 1\} \subset X$ may be regarded as a unit ball in some Hilbert space $X^G_\varphi \subset X$ with norm $\|x\|_\varphi := \|\varphi(G)^{-1}x\|$. In case of a power function $\varphi(t) = t^p$ we shall abbreviate $\|x\|_p := \|x\|_{t^p}$. The collection

$$\left\{ X^G_\varphi, \quad \varphi \text{ is an index function} \right\}$$

constitutes a variable Hilbert scale, similar to the one introduced in [9, 10].

3.1. Interpolation inequality for elements.

Here we recall the variant as outlined in [17].

Theorem 3.1. *Let* φ, ψ *be proper index functions. If* $t \mapsto \varphi^2\left(\left(\psi^2\right)^{-1}(t)\right)$ *is concave, then for any index function* θ *and any* $0 \neq v \in X_\theta^G$ *we have*

$$\varphi^{-1}\left(\frac{\|v\|_{\theta/\varphi}}{\|v\|_\theta}\right) \leq \psi^{-1}\left(\frac{\|v\|_{\theta/\psi}}{\|v\|_\theta}\right). \tag{3.1}$$

As an immediate application we mention the following upper bound for the modulus of continuity from (2.4), which indicates that in general the lower bound from Theorem 2.1 is tight.

Corollary 3.1. *Let Assumption A.1 hold true. If the function* $t \mapsto \left[(\Psi_p^2)^{-1}(t)\right]^{2p}$ *is concave then*

$$\omega(G_{p,E}, \delta) \leq E\left[(\Psi_p^{-1}(\delta/mE)\right]^p. \tag{3.2}$$

Proof. By using the link assumption A.1 we can bound

$$\omega(G_{p,E}, \delta) \leq \sup\left\{\|x\|, \quad \|x\|_p \leq E,\ \|x\|_{1/\varrho} \leq \delta/m\right\}.$$

Thus, we let $\theta(t) = \varphi(t) = t^p$ and $\psi(t) := \Psi_p(t)$ and apply inequality (3.1), which gives (3.2). □

The interpolation inequality (3.1) can be used in various other cases, see e.g. [17]. However, its application is limited to bounding norms of individual elements in Hilbert space; the application to bounding norms of operators is limited, but this also has been explored in [17, Appendix A].

3.2. Interpolation inequality for operators.

Interpolation of linear operators acting between Hilbert spaces will be a main tool in our subsequent analysis and we shall establish a variant, which perfectly fits our applications, but it may be of independent interest. It has its origin in the seminal paper [4], related to Theorem 3.3, below.

Interpolation of operators between Hilbert scales is based on the following partial ordering for self-adjoint operators. Let G and H be self-adjoint operators in some Hilbert spaces X. We say that $G \leq H$ if for all $x \in \mathcal{D}(H)$ the inequality $\langle Gx, x\rangle \leq \langle Hx, x\rangle$ holds true.

Definition 1. Let $f\colon [0,a] \to \mathbb{R}^+$ be a continuous function.

(i) It is called *operator monotone* if for any pair $G, H \geq 0$ of self-adjoint operators with spectra in $[0,a]$ the ordering $G \leq H$ implies $f(G) \leq f(H)$.
(ii) It is called *operator concave* if for any pair $G, H \geq 0$ of self-adjoint operators with spectra in $[0,a]$ we have

$$f(\frac{G+H}{2}) \geq \frac{f(G)+f(H)}{2}. \tag{3.3}$$

Remark 2. A detailed analysis of operator monotone, operator concave functions can be found in [5], or the more recent [2]. In particular we mention that any non-negative operator concave function on some interval is necessarily operator monotone ([2, Proof of Thm. V.2.5]). Also, any non-negative function, which is operator monotone on $[0, \infty)$ is operator concave. Moreover, if f, g are a pair of non-negative operator concave functions for which the composition $t \mapsto f(g(t))$ is defined, then this composition is also non-negative operator concave.

We shall use the following characterization.

Theorem 3.2 ([2, Thm. V.2.3] and [8])**.** *A function f from above is operator concave if and only if for every contraction K the inequality*

$$K^* f(G) K \leq f(K^* G K) \quad \textit{holds true.} \tag{3.4}$$

Example 1. Of special importance are the functions $t \mapsto t^\theta,\ 0 < \theta \leq 1$, which are known to be both, operator monotone and concave on $[0, \infty)$.

In the context of ill-posed problems the functions $t \mapsto \log^{-p}(1/t),\ 0 < t < 1$ gained importance. These functions are operator concave provided $0 < p \leq 1$.

Finally we mention that the function $t \to -1/t$ is operator monotone, hence any estimate $\|Ax\| \leq \|Bx\|$ implies $\|B^{-1}x\| \leq \|A^{-1}x\|,\ x \in \mathcal{D}(A^{-1})$.

The following result is basic for further considerations. In a similar form this was established in [4]. A further historical account is given there. The proof which is given below is based on Theorem 3.2.

Theorem 3.3. *Let $f^2 \colon [0, a] \to \mathbb{R}^+$ be an operator concave proper index function and $S \colon X \to Y$ be a bounded operator. For every pair G, H of operators, for which $\max\{\|G^*G\|, \|H^*H\|\} \leq a$ the following assertion holds true: If there are constants $C_1, C_2 < \infty$ such that*

$$\|Sx\| \leq C_1 \|x\|, \quad x \in X, \tag{3.5}$$

and

$$\|GSx\| \leq C_2 \|Hx\|, \quad x \in X, \tag{3.6}$$

then

$$\|f(G^*G)Sx\| \leq \max\{C_1, C_2\} \, \|f(H^*H)x\|, \quad x \in X. \tag{3.7}$$

Proof. We rewrite the assumptions (3.5) and (3.6) as

$$\frac{S^*S}{C_1^2} \leq \mathrm{I} \quad \text{and} \quad \frac{S^*G^*GS}{C_2^2} \leq H^*H$$

and distinguish two cases. If $C_2 \leq C_1$ then

$$\frac{S^*G^*GS}{C_1^2} \leq \frac{S^*G^*GS}{C_2^2} \leq H^*H.$$

Theorem 3.2 and operator concavity imply

$$\frac{S^*}{C_1} f^2(G^*G) \frac{S}{C_1} \leq f^2(\frac{S^*G^*GS}{C_1^2}) \leq f^2(H^*H),$$

hence (3.7).

Otherwise $C_1 \leq C_2$ and we rewrite

$$\frac{S^*G^*GS}{C_2^2} = \frac{S^*}{C_1}\left(\frac{C_1^2}{C_2^2}G^*G\right)\frac{S}{C_1}.$$

Now we use the fact, see Lemma 3.1 below, that for concave proper index functions f^2 and for $\kappa > 1$ we have $f^2(\kappa t) \leq \kappa f^2(t)$, whenever $0 < \kappa t \leq a$, thus

$$\frac{S^*}{C_1}f^2(G^*G)\frac{S}{C_1} = \frac{S^*}{C_1}f^2(\frac{C_2^2}{C_1^2}\frac{C_1^2}{C_2^2}G^*G)\frac{S}{C_1} \leq \frac{C_2^2}{C_1^2}\frac{S^*}{C_1}f^2(\frac{C_1^2}{C_2^2}G^*G)\frac{S}{C_1}. \tag{3.8}$$

We arrive, using Theorem 3.2 again, at

$$\frac{S^*}{C_1}f^2(G^*G)\frac{S}{C_1} \leq \frac{C_2^2}{C_1^2}f^2(\frac{S^*}{C_1}\frac{C_1^2}{C_2^2}G^*G\frac{S}{C_1}) \leq \frac{C_2^2}{C_1^2}f^2(H^*H),$$

from which (3.7) is easily obtained. □

Remark 3. It is worth discussing the case when the function f is of power type, say $f(t) := t^{\theta/2}$, with $0 < \theta \leq 1$ to have f^2 concave. In this case the estimate from above refines the bound (3.7) to $\|f(G^*G)Sx\| \leq C_1^{1-\theta}C_2^{\theta}\|f(H^*H)x\|, \quad x \in X$.

The following result extends Theorem 3.3 to the present setup.

Theorem 3.4. *Let $G, H \geq 0$ be self-adjoint operators with spectra in $[0, b]$ and $[0, a]$, respectively. Furthermore, let φ, ρ and r be proper index functions on intervals $[0, b]$ and $[0, a]$, respectively, such that $b \geq \|G\|$ and $\rho(b) \geq r(a)$. Then the function*

$$f(t) := \varphi((\rho^{-1})(r(t))), \quad 0 < t \leq a,$$

is well defined. The following assertion holds true: If $t \to \varphi^2((\rho^2)^{-1}(t))$ is operator concave on $[0, \rho^2(b)]$ [1], then

$$\|Sx\| \leq C_1\|x\|, \quad x \in X, \tag{3.9}$$

and

$$\|\rho(G)Sx\| \leq C_2\|r(H)x\|, \quad x \in X, \tag{3.10}$$

yield

$$\|\varphi(G)Sx\| \leq \max\{C_1, C_2\}\|f(H)x\|, \quad x \in X. \tag{3.11}$$

Proof. By definition of f we obtain

$$f^2((r^2)^{-1}(t)) = \varphi^2((\rho^2)^{-1}(t)), \quad 0 < t \leq r^2(a).$$

Now let $\tilde{H} := r(H)$ and $\tilde{G} := \rho(G)$. By assumption both operators have spectrum in $(0, \rho(b)]$, and moreover $\|\tilde{G}Sx\| \leq C_2\|\tilde{H}x\|$, such that Theorem 3.3 is applicable and implies

$$\|\varphi((\rho^2)^{-1}(\tilde{G}^*\tilde{G}))Sx\| \leq \max\{C_1, C_2\}\|f((r^2)^{-1}(\tilde{H}^*\tilde{H}))x\|,$$

from which the proof can easily be completed. □

---

[1] This means that it has an operator concave extension from $(0, \rho^2(b)]$ to $[0, \rho^2(b)]$.

$$\begin{tikzcd}
G : X_\rho^G \arrow[r, "J_{\rho,G}^{\varphi,G}"] \arrow[d, "S"] & X_\varphi^G \arrow[r, "J_{\varphi,G}"] \arrow[d, "S"] & X \arrow[d, "S"] \\
H : Y_r^H \arrow[r, "J_{r,H}^{f,H}"] & Y_f^H \arrow[r, "J_{f,H}"] & Y
\end{tikzcd}$$

Figure 1. Then setup of interpolation

The above results have an important translation to variable Hilbert scales and assert that the pairs $(X, X_{1/r})$ and $(Y, Y_{1/\rho})$, and by duality (X, X_r) and (Y, Y_ρ), are actually *exact interpolation pairs*, extending an analogous result from [4]. In these terms Theorem 3.4 admits a different, more suggestive interpretation.

Corollary 3.2. *Let the scales* $\left\{X_\varphi^G\right\}$ *and* $\left\{Y_f^H\right\}$ *be variable Hilbert scales, generated by non-negative operators* G *and* H *with* $a := \|H\|$. *Let* φ, ρ *on* $(0, b]$ *be proper index functions such that* $b \geq \|G\|$ *and* $t \mapsto \varphi^2((\rho^2)^{-1}(t))$ *is operator concave on* $[0, \rho^2(b)]$. *Given another proper index function* r *on* $(0, a]$ *with* $r(a) \leq \rho(b)$ *we assign*

$$f(t) := \varphi((\rho^{-1})(r(t))), \quad 0 < t \leq a.$$

If $S\colon X \to Y$ *is bounded, both as* $S\colon X \to Y$ *and* $S\colon X_\rho^G \to Y_r^H$, *then* S *acts boundedly from* X_φ^G *to* Y_f^H *and*

$$\|S\colon X_\varphi^G \to Y_f^H\| \leq \max\left\{\|S\colon X \to Y\|, \|S\colon X_\rho^G \to Y_r^H\|\right\}. \tag{3.12}$$

Proof. Let us denote $T := S^*$. The assumption $\|S\colon X_\rho^G \to Y_r^H\| \leq C_2$ is equivalent to $\|T\colon Y_{1/r}^H \to X_{1/\rho}^G\| \leq C_2$, which in turn translates to

$$\|\rho(G)Tv\| \leq C_2\|r(H)v\|, \ v \in X.$$

Similarly, $\|T\colon Y \to X\| \leq C_1$. By Theorem 3.4 we obtain

$$\|\varphi(G)Tv\| \leq \max\{C_1, C_2\}\,\|f(H)v\|, \ v \in X,$$

which implies $\|S\colon X_\varphi^G \to Y_f^H\| \leq \max\{C_1, C_2\}$. □

Remark 4. Again the power case $f(t) = t^{\theta/2}$ is worth mentioning. As discussed in Remark 3, we are able to refine the bound from (3.12) to

$$\|S\colon X_\varphi^G \to Y_f^H\| \leq \|S\colon X \to Y\|^{1-\theta}\|S\colon X_\rho^G \to Y_r^H\|^\theta. \tag{3.13}$$

Interpolation with such bound is said to be of *type* θ, see [1, Chapt. 2.4]. It is not clear to the authors whether there is a sensible generalization for general operator concave functions, or whether this is an artifact for power functions.

The situation as described in the theorem can be visualized as in Figure 1, with $J_{\rho,G}^{\varphi,G}, J_{\varphi,G}$ and $J_{r,H}^{f,H}, J_{f,H}$ describing the canonical embeddings. In our notation we shall drop the involved operators G and H when this is clear from the context. The position of X_φ^G between X_ρ^G and X on top is given by the function $t \to \varphi((\rho)^{-1}(t))$ and f is determined in such a way that Y_f^H has the appropriate position in the scale on bottom.

Remark 5. Note that the roles of G and H are completely interchangeable. Thus if T is boundedly invertible, both between X and Y and between X_ρ^G and Y_r^H, then so is $T\colon X_\varphi^G \to Y_f^H$, with a corresponding norm bound.

3.3. **Interpolation for embeddings.** Interpolation in the form of Corollary 3.2 will be used only for operators S which are multiples of the identity mapping. For such operators we can extend our analysis from operator concave to operator monotone functions. To this end let us return to the context of Theorem 3.3. If there the mapping $S := C_1\,\mathrm{I}$ and $C_2 \leq C_1$ then $C = \max\{C_1, C_2\} = C_1$ and estimate (3.8) implies (3.7). Otherwise we impose the following assumption on a non-negative function $f\colon (0,a] \to \mathbb{R}^+$.

Assumption A.3. There is a constant D_f for which

$$f(t)/t \leq D_f f(u)/u, \qquad \text{whenever } 0 < u < t \leq a.$$

This assumption holds true in many cases and we list two important ones.

Lemma 3.1. *Assume that f is a concave proper index function on $(0,a]$. Then*

$$f(t)/t \leq f(u)/u, \quad \text{whenever} \quad 0 < u \leq t \leq a, \tag{3.14}$$

Proof. Let $0 < u \leq t \leq a$, hence we can convexly combine $u = (u/t)t + (1-u/t)0$ with $0 < u/t \leq 1$. Since f is assumed to be a proper and concave index function we conclude that

$$f(u) \geq \frac{u}{t} f(t) + \left(1 - \frac{u}{t}\right) f(0) \geq \frac{u}{t} f(t),$$

from which the proof can be completed. □

Remark 6. We note that in this case we may let $D_f = 1$. This holds true for the classical context of powers $t \mapsto t^\theta$, with $0 < \theta \leq 1$, but also for logarithmic functions $t \mapsto \log^{-\mu}(1/t)$ with $0 < \mu \leq 1$ as these are met for severely ill-posed problems.

Furthermore we recall the following result from [19, Lemma 3].

Lemma 3.2. *Let $f\colon [0,c']$ be non-negative operator monotone and $0 < a < c'$. There is $D_f = D_f(c'-a) \geq 1$ such that*

$$f(t)/t \leq D_f f(u)/u, \qquad \text{whenever } 0 < u < t \leq a.$$

We state the following consequence.

Corollary 3.3. *Let $X = Y$ and $S = C_1\,\mathrm{I}$ be a multiple of the identity. Under the assumptions of Theorem 3.3 the following assertions hold true.*

If $C_2 \leq C_1$ and the function f^2 is operator monotone on $(0,a]$ then the estimate (3.7) holds true.

If $C_1 < C_2$ and f^2 is an operator monotone proper index function which obeys Assumption A.3, then the estimate (3.7) holds true with the additional factor $\sqrt{D_{f^2}}$.

Proof. The first assertion was already discussed. To prove the second one we start with the following observation. For every $c > 1$ Assumption A.3 implies that

$$f^2(t) \le D_{f^2}\ cf^2(t/c), \quad 0 < t \le a,$$

by letting $u := t/c$. Spectral calculus allows to extend this to arbitrary self-adjoint operators with spectra in $(0, a]$, in particular we have

$$f^2(G^*G) \le D_{f^2}\ cf^2(G^*G/c). \tag{3.15}$$

Thus, if $S = C_1 \mathrm{I}$, then (3.6) yields $G^*G \le C_2^2/C_1^2 H^*H$. If $C_2 \le C_1$ then operator monotonicity yields $f^2(G^*G) \le f^2(H^*H)$, which implies (3.7). Otherwise, if we let $c := C_2^2/C_1^2 > 1$, then estimate (3.15) implies

$$C_1^2 f^2(G^*G) \le C_1^2 D_{f^2} c f^2(G^*G/c) \le C_1^2 D_{f^2} c f^2(H^*H) = D_{f^2} C_2^2 f^2(H^*H),$$

which allows to complete the proof in this case. □

Analogously this can be extended to the situation in Corollary 3.2 and we state the following analog.

Corollary 3.4. *Let $X = Y$ and $S := \mathrm{I}$ be the identity. Under the assumptions of Corollary 3.2 the following assertions hold true.*

If $C_2 \le 1$ and the function f^2 is operator monotone on $(0, a]$ then

$$\| J^{f,H}_{\varphi,G} \colon X^G_\varphi \to X^H_f \| \le 1. \tag{3.16}$$

If $C_2 > 1$ and f^2 is an operator monotone proper index function which obeys Assumption A.3, then

$$\| J^{f,H}_{\varphi,G} \colon X^G_\varphi \to X^H_f \| \le \sqrt{D_{f^2}} C_2. \tag{3.17}$$

4. LINEAR INVERSE PROBLEMS: PRELIMINARY ANALYSIS

Let us consider a general regularization scheme, say $y^\delta \mapsto g_\alpha(T^*T)T^*y^\delta$ in Hilbert spaces, related to some bounded operator $T \colon X \to Y$. Of course, the choice of operator T cannot be arbitrary and must be related to the underlying operator A from equation (1.1). We introduce the related (pure) residual as

$$r_\alpha(t) := 1 - tg_\alpha(t), \quad t \in (0, \|T\|^2],$$

and recall the following definition from [18, Def. 2].

Definition 2. A family $g_\alpha(t)$, $\alpha > 0$, of piece-wise continuous functions is called *regularization*, if there exist positive constants $\gamma \ge 1$ and γ_* such that

$$\sup_{0<t\le\|T\|^2} t^{1/2} |g_\alpha(t)| \le \gamma_*/\sqrt{\alpha}, \tag{4.1}$$

and

$$\sup_{0<t\le\|T\|^2} |r_\alpha(t)| \le \gamma. \tag{4.2}$$

In particular there is a constant $\beta < \infty$ for which

$$\sup_{0<t\le\|T\|^2} t\,|g_\alpha(t)| \le \beta. \tag{4.3}$$

Requirement (4.1) is the usual normalization, we shall dwell on it in more detail in Section 4.4. Requirement (4.2) is necessary for convergence on exact data $y = Ax^\dagger$ if we let $T := A$. Furthermore we may always find $\beta \le 1 + \gamma$. For example, Tikhonov regularization has $\gamma = \beta = 1$.

This is not enough and we assume the following qualification, as this was introduced in [18].

Definition 3. Let φ be a proper index function. The regularization g_α is said to have qualification φ with constant γ if

$$\sup_{0<t\le\|T\|^2} |r_\alpha(t)|\,\varphi(t) \le \gamma\varphi(\alpha), \quad 0 < \alpha \le \|T\|^2. \tag{4.4}$$

This generalizes the notion of qualification as it was given in [32] to the case of smoothness given in terms of general source conditions. The standard regularization methods in Hilbert scales such as

(a) Tikhonov regularization with $g_\alpha(t) = 1/(t+\alpha)$,
(b) Asymptotical regularization with $g_\alpha(t) = \frac{1}{t}(1 - e^{-t/\alpha})$,
(c) Landweber iteration with $g_\alpha(t) = \frac{1}{t}\left(1 - (1-t)^{1/\alpha}\right)$ or,
(d) Spectral method with $g_\alpha(t) = 1/t$ for $t \ge \alpha$ and $g_\alpha(t) = 0$ for $t < \alpha$,

satisfy assumptions (4.1)– (4.3) with constants less than or equal to one.

4.1. **Impact of Assumption A.3 on the qualification.**

Our analysis will use the following sufficient condition for an index function to be a qualification of the regularization g_α.

Lemma 4.1. *Suppose g_α is a regularization with qualification φ and constant γ. If $\psi(t)$ is any proper index function such that for some $D < \infty$ it holds true that*

$$\frac{\psi(t)}{\varphi(t)} \le D\frac{\psi(u)}{\varphi(u)}, \quad \text{whenever} \quad 0 < u \le t \le a, \tag{4.5}$$

then g_α has qualification ψ with constant $D\gamma$.

Proof. To bound $|r_\alpha(t)|\,\psi(t)$ we shall distinguish two cases. For $0 < t \le \alpha$ the required bound follows from monotonicity. Otherwise we conclude

$$|r_\alpha(t)|\,\psi(t) = |r_\alpha(t)|\,\varphi(t)\frac{\psi(t)}{\varphi(t)} \le \gamma\varphi(\alpha)\sup_{t\ge\alpha}\frac{\psi(t)}{\varphi(t)} \le \gamma\varphi(\alpha)D\frac{\psi(\alpha)}{\varphi(\alpha)},$$

from which the proof can be completed. □

We shall draw conclusions for the qualification under the following standing assumption on the regularization.

Assumption A.4. For γ as above we assume that

$$\sup_{0<t\leq\|T\|^2} |r_\alpha(t)|\, t^{1/2} \leq \gamma\sqrt{\alpha}, \tag{A.4.1}$$

or the stronger

$$\sup_{0<t\leq\|T\|^2} |r_\alpha(t)|\, t \leq \gamma\alpha. \tag{A.4.2}$$

Remark 7. Assumption A.4.1 says that the qualification is at least $t \mapsto \sqrt{t}$, which is a very weak requirement, and which is fulfilled for most regularization schemes, in particular for the regularizations from (a)–(d) above. We stress that A.4.2 is stronger than A.4.1, still this holds for most regularizations, in particular those from above.

Altogether, Lemmas 4.1 and 3.1 allow us to draw the following conclusions about the qualification of g_α.

Corollary 4.1. *Suppose the proper index function g^2 obeys A.3 on $[0, \|T^2\|]$.*

(i) *If the regularization g_α obeys Assumption A.4.1 with constant γ then g_α has qualification g with constant $\sqrt{D_{g^2}}\gamma$.*
(ii) *If the regularization g_α obeys Assumption A.4.2 with constant γ then g_α has qualification $g(t)\sqrt{t}$ with constant $\sqrt{D_{g^2}}\gamma$.*

4.2. **Regularization with additional smoothing.** In our analysis we shall deal with regularization based on the operator $T := AG^s$, in which the regularized approximations with exact and noisy data y and y^δ, respectively, are defined by

$$x_\alpha := G^s g_\alpha(T^*T)T^*y \quad \text{and} \quad x_\alpha^\delta := G^s g_\alpha(T^*T)T^*y^\delta. \tag{4.6}$$

Here, g_α is a regularization and the parameter s controls the smoothness which is introduced into (or removed from) the regularization process. We refer to the discussion in Remark 10. For deriving order optimal error bounds for $\|x_\alpha^\delta - x^\dagger\|$ with x_α^δ defined by (4.6) the following error representations are useful and will be exploited at different places. We express the different components of the error as

$$x_\alpha^\delta - x_\alpha = G^s g_\alpha(T^*T)T^*(y^\delta - y), \tag{4.7}$$

$$x^\dagger - x_\alpha = G^s r_\alpha(T^*T)G^{-s}x^\dagger. \tag{4.8}$$

In our analysis below we shall use the convention that spaces X_φ^G are generated by the operator G, whereas the respective spaces X_f^H generated by $H := m^{-2}T^*T$, with m from A.1.

4.3. **Impact of operator monotonicity on norm bounds.** Our analysis will use factorization of operators through different spaces in Hilbert scales. Given a parameter s, we recall $\Psi_s(t) := t^s\varrho(t),\ 0 < t \leq b$ from (2.5). We assume that s is sufficiently large such that Ψ_s is a proper index function. The initial point is Assumption A.1, which, using the present notation can be rewritten as $\|\Psi_s(G)x\| \leq \|H^{1/2}x\|$, or equivalently as

$$\|J_{\Psi_s}^{1/2} \colon X_{\Psi_s}^G \to X_{t^{1/2}}^H\| \leq 1. \tag{4.9}$$

$$\begin{tikzcd}
G : X^G_{\Psi_s} \arrow[r, "J^{p-s,G}_{\Psi_s,G}"] \arrow[d, "J^{1/2,H}_{\Psi_s,G}"] & X^G_{t^{p-s}} \arrow[r, "J_{p-s,G}"] \arrow[d, "J^{g,H}_{p-s,G}"] & X \arrow[r, "J^{1/\Psi_s,G}"] \arrow[d, "J"] & X^G_{1/\Psi_s} \arrow[d, "J^{-1/2,H}_{1/\Psi_s,G}"] \\
H : X^H_{\sqrt{t}} \arrow[r, "J^{g,H}_{1/2,H}"] & X^H_g \arrow[r, "J_{g,H}"] & X \arrow[r, "J^{-1/2,H}"] & X^H_{1/\sqrt{t}}
\end{tikzcd}$$

FIGURE 2. Description of the setup

Analogously we replace Assumption A.2 by

$$\|J^{-1/2}_{1/\Psi_s} \colon X^G_{1/\Psi_s} \to X^H_{1/t^{1/2}}\| \leq M/m. \tag{4.10}$$

We start with the consequences of the bounds (4.9) and (4.10) in the light of interpolation. To this end let us introduce the function

$$g(t) := \left[\Psi_s^{-1}(\sqrt{t})\right]^{p-s}, \quad 0 < t \leq \Psi_s^2(b), \tag{4.11}$$

with function Ψ_s from (2.5). Note that for $s \leq p$ the function g is non-decreasing, whereas for $s > p$ this is to hold true for $1/g$, such that in our analysis we are to distinguish these cases, called the high order and low order case, respectively. In any case, the following result holds true.

Lemma 4.2. *The function $t \mapsto g(t)\sqrt{t}$, $0 < t \leq \Psi_s^2(b)$ is a proper index function.*

Proof. Since ϱ is supposed to be a proper index function, the function

$$u \mapsto \left(\Psi_p^2\right)^{-1}(u), \ 0 < u \leq b$$

is increasing. Thus substituting $u := \left(\Psi_s^2\right)^{-1}(t)$ we obtain

$$tg^2(t) = \Psi_s^2(u)u^{2(p-s)} = \Psi_p^2(u), \quad 0 < u \leq b,$$

which is increasing, as mentioned above. This allows to complete the proof. □

Because we shall frequently use norm bounds as in (4.9) or (4.10), we make the following assumption, covering what is needed to apply both Corollary 4.1 and Proposition 4.1, below. For simplicity we shall confine our presentation to the case when the constant D_f in Assumption A.3 equals one. The general case is easily obtained analogously.

Assumption A.5. Throughout we assume that the parameter s is chosen such that Ψ_s is increasing. The function g is defined on some interval $[0, \Psi_s^2(b)]$ for some b with $\Psi_s^2(b) \geq a$.

- **$\mathbf{p} \leq \mathbf{s}$:** The function $t \mapsto 1/g^2(t)$ is operator monotone.
- **$\mathbf{s} < \mathbf{p}$:** The function $t \mapsto g^2(t)$ is operator monotone and obeys Assumption A.3 with constant $D_{g^2} = 1$, hence the function $t \mapsto t/g^2(t)$ is increasing.

Remark 8. By virtue of Assumption A.5, necessarily the inclusions $X^H_{\sqrt{t}} \subseteq X^H_g \subseteq X^H_{1/\sqrt{t}}$ hold true. The latter can also be seen from Lemma 4.2. Figure 2 visualizes the setup, where we exhibit the low order case ($s < p$).

By (4.9) and (4.10), the assumptions A.1 and A.2 ensure the boundedness of the embeddings on the left and right, respectively.

Further note that in the low order case this imposes restrictions on the (possible) negativity of the parameter s.

Proposition 4.1. *Let the function g obey Assumption A.5.*

$\mathbf{p \le s}$:
  - (i) *Under A.2 we have* $\|J^g_{p-s}\colon X^G_{t^{p-s}} \to X^H_g\| \le M/m$,
  - (ii) *Under A.1 we have* $\|J^{1/g}_{s-p}\colon X^G_{1/t^{p-s}} \to X^H_{1/g}\| \le 1$.

$\mathbf{s < p}$:
  - (i) *Under A.1 we have* $\|J^g_{p-s}\colon X^G_{t^{p-s}} \to X^H_g\| \le 1$,
  - (ii) *Under A.2 we have* $\|J^{1/g}_{s-p}\colon X^G_{1/t^{p-s}} \to X^H_{1/g}\| \le M/m$.

Proof. We only prove the first assertion, the proofs of the other assertions are similar. Under A.2 and by definition of H we derive from (4.10) that

$$\|(J^{-1/2}_{1/\Psi_s})^*\colon X^H_{t^{1/2}} \to X^G_{\Psi_s}\| \le \frac{M}{m}.$$

Thus we apply Corollary 3.4, in the light of Remark 5, with $r(t) := \sqrt{t}$, $\rho(t) := \Psi_s(t)$ and $\varphi(t) := 1/g(t)$. We can complete the proof since $(1/g)^2(t) = \varphi^2((\rho^2)^{-1}(t))$ is supposed to be operator monotone. □

4.4. **Controlling the noise in different norms.** In the traditional concept of regularization in Hilbert scales the normalization (4.1) is used to bound the noise amplification in the target space X. In operator terms this may be rewritten as

$$\|g_\alpha(T^*T)\colon X^H_{\sqrt{t}} \to X\| \le \frac{\gamma_*}{m\sqrt{\alpha}} \quad \left(= \frac{\gamma_*}{m^2}\frac{1}{\sqrt{\alpha/m^2}}\right), \tag{4.12}$$

because the scale is generated by $H = T^*T/m^2$. In our subsequent analysis we shall have to control norm bounds in target spaces different from X, in particular in X^H_g. Under the Assumption A.5 of operator monotonicity and as outlined in Remark 8, the extremal cases obtained in this way are when $g(t) = \sqrt{t}$ and $g(t) = 1/\sqrt{t}$.

The particular case $g(t) = 1/\sqrt{t}$ is covered by property (4.3).

Lemma 4.3. *For any regularization g_α it holds true that*

$$\|g_\alpha(T^*T)\colon X^H_{\sqrt{t}} \to X^H_{1/\sqrt{t}}\| \le \frac{\beta}{m^2}. \tag{4.13}$$

Proof. Using (4.3) we obtain

$$\|g_\alpha(T^*T)\colon X^H_{\sqrt{t}} \to X^H_{1/\sqrt{t}}\| = \|\frac{1}{m^2}T^*Tg_\alpha(T^*T)\| = \frac{1}{m^2}\sup_{0<t\le a}|tg_\alpha(t)| \le \frac{\beta}{m^2},$$

and the proof is complete. □

We need to control the noise amplification in the target space $X^H_{\sqrt{t}}$. Assumptions which have to be made *must* be consistent with (4.1) and (4.13). This may be viewed as an extrapolation problem and leads to the following requirement for a regularization g_α.

Assumption A.6. There is a constant γ^* such that

$$|g_\alpha(t)| \le \frac{\gamma^*}{\alpha} \quad \left(= \frac{\gamma^*}{m^2}\frac{1}{\alpha/m^2}\right). \tag{4.14}$$

This assumption is indeed consistent with (4.1) as shown in [32, Chapt. 2.3]. We shall state the consequences of this normalization.

Proposition 4.2. *Under the assumption that g obeys A.5 and that the regularization obeys A.6 the following bounds hold true.*

$$\|g_\alpha(T^*T)\colon X^H_{\sqrt{t}} \to X^H_g\| \le \frac{\max\{\beta, \gamma^*\}}{m^2}\frac{1}{\sqrt{\alpha/m^2}g(\alpha/m^2)}. \tag{4.15}$$

Proof. We start as usual with the representation

$$\|g_\alpha(T^*T)\colon X^H_{\sqrt{t}} \to X^H_g\| = \sup_{0<t\le a} |g_\alpha(t)| \frac{\sqrt{t/m^2}}{g(t/m^2)}.$$

Under A.5, in the low order case $0 \le s < p$, the function g^2 is operator monotone, in particular it is non-decreasing, such that $\sqrt{t}g(t)$ is increasing. Moreover, Assumption A.3 applies to g^2 with constant one. Thus for $0 < t \le \alpha$ we conclude that

$$|g_\alpha(t)| \frac{\sqrt{t/m^2}}{g(t/m^2)} \le \frac{\gamma^*}{\alpha}\frac{\sqrt{t/m^2}}{g(t/m^2)} \le \frac{\gamma^*}{m^2}\frac{1}{\alpha/m^2}\frac{\sqrt{\alpha/m^2}}{g(\alpha/m^2)} = \frac{\gamma^*}{m^2}\frac{1}{\sqrt{\alpha/m^2}g(\alpha/m^2)}.$$

Otherwise $t > \alpha$ and we conclude, using that $s \mapsto 1/(\sqrt{s}g(s))$ is decreasing,

$$|g_\alpha(t)| \frac{\sqrt{t/m^2}}{g(t/m^2)} = |g_\alpha(t)|\, t\frac{1}{m^2}\frac{1}{\sqrt{t/m^2}g(t/m^2)} \le \frac{\beta}{m^2}\frac{1}{\sqrt{\alpha/m^2}g(\alpha/m^2)},$$

which completes the proof in the low order case. In the high order case, a similar reasoning allows to draw similar conclusions. □

4.5. **Examples.**

Finitely smoothing case. Let us assume that the operators A^*A and G are related by

$$A^*A = G^{2a} \tag{4.16}$$

where a is some positive constant. In this case both assumptions A.1 and A.2 hold true as equality with $\varrho(t) = t^a$, $m = 1$ and $M = 1$. We easily see that the function ϱ is a proper index function and that the function Ψ_s from (2.5) attains the form

$\Psi_s(t) = t^{a+s}$. Since $\Psi_s^{-1}(\sqrt{t}) = t^{1/(2a+2s)}$ we obtain that the function g as defined in (4.11) possess the representation

$$g(t) = t^{\frac{p-s}{2(a+s)}}.$$

Since by [2, Thm. V.1.9] power functions t^ν are operator concave on $[0,\infty)$ for $0 \leq \nu \leq 1$, we conclude that under the natural constraints that $p \geq 0$, $a > 0$ and $s > -a$ the following statements are true:

(i) g^2 is an operator concave function for $s \leq p \leq 2s + a$.
(ii) $1/g^2$ is an operator concave function for $p \leq s$.

Thus Assumption A.5 is satisfied, provided that $s \geq (p-a)/2$, which limits the negativity of the parameter s.

Infinitely smoothing case. Here A^*A and G are related by

$$A^*A = e^{-\frac{1}{2}G^{-\mu}} \quad \text{for some } \mu > 0.$$

This corresponds to regularization of certain *severely ill-posed problems*, in particular the *backward heat equation*, where this holds true for $G := (-\Delta)^{-1/2}$ and $\mu = 2$. In this special case both assumptions A.1 and A.2 hold true as equality with $\varrho(t) = e^{-\frac{1}{4}t^{-\mu}}$. Below we shall use results from [29]. The inverse function of $\varrho^2(t) = e^{-\frac{1}{2}t^{-\mu}}$ is obtained as

$$(\varrho^2)^{-1}(s) = 2^{-1/\mu}\log^{-1/\mu} 1/s, \quad 0 < s < 1.$$

Note that both the functions $u \mapsto -1/u$ and $u \mapsto \log u$ are operator concave on $(0,\infty)$ and so is their composition, such that the function $u \mapsto (\varrho^2)^{-1}(u),\ 0 < u < 1$ is operator concave for $\mu \geq 1$. Following [31, Lem. 2.4] the function $\Psi_s^2(t) = t^{2s}\varrho^2(t), \quad t > 0$, has an operator monotone inverse on $[0,\infty)$, whenever $0 \leq s \leq 1/2$ and $\mu \geq 1$. In the light of Remark 2 it is thus operator concave, and this extends to the composition with the non-negative operator concave function $t \mapsto t^\theta$, $(0 < \theta \leq 1)$. Therefore, applying again [31, Lem. 2.4] the function

(i) g^2 related to (4.11) is operator concave provided $s \leq p \leq 2s$,
(ii) $1/g^2$ is operator concave for $0 < p \leq s$.

Summarizing, the assumptions of operator concavity are fulfilled for $0 \leq s \leq 1/2$ and $\mu \geq 1$. Therefore, if $\|T\| < 1$, then Assumption A.5 holds.

We emphasize that high order regularization with $s > 1/2$ for severely ill-posed problems is not covered in the present approach. The case $s = 0$ was treated in [3, Example 4.3].

5. Bounding the error in different norms

We turn to analyzing the error terms and start with the following observation, which is easy to verify and we omit the proof. For any pair of index functions f and g and operator $T\colon X \to X$ it holds true that

$$\|T\colon X_f^H \to X_g^H\| = \|(1/g)(H)Tf(H)\colon X \to X\|. \tag{5.1}$$

Below we shall analyze the error in different norms. On the one hand this provides insight into different requirements to be imposed to achieve norm bounds but on the other hand this will allow us to finally bound the error $\|x_\alpha^\delta - x^\dagger\|_X$ in the original norm.

To this end let λ be any index function, such that $X_{t^p}^G \subset X_\lambda^G$. Then the representations (4.7) and (4.8) together with the bias variance decomposition

$$\|x_\alpha^\delta - x^\dagger\|_\lambda \le \|x_\alpha^\delta - x_\alpha\|_\lambda + \|x_\alpha - x^\dagger\|_\lambda, \tag{5.2}$$

yield, using (4.7), the estimate

$$\|x_\alpha^\delta - x_\alpha\|_\lambda = \|(1/\lambda)(G)G^s g_\alpha(T^*T)T^*(y^\delta - y)\| \le \delta m \|g_\alpha(T^*T)\colon X_{\sqrt{t}}^H \to X_{\lambda(t)t^{-s}}^G\|.$$

Similarly, (4.8) yields

$$\|x^\dagger - x_\alpha\|_\lambda = \|(1/\lambda)(G)G^s r_\alpha(T^*T)G^{p-s}v\| \le E\|r_\alpha(T^*T)\colon X_{t^{p-s}}^G \to X_{\lambda(t)t^{-s}}^G\|.$$

We shall bound the right hand sides based on the following factorizations through different spaces. Given any functions ϕ and ψ, let $J_\phi^\psi\colon X_\phi^G \to X_\psi^H$ denote the canonical embeddings (possibly unbounded). With this notation we have, with the function l yet to be determined,

$$\begin{tikzcd} g_\alpha(T^*T) : X_{\sqrt{t}}^H \arrow[r, "g_\alpha(T^*T)"] & X_l^H \arrow[r, "(J^{1/l}_{1/\lambda(t)t^{-s}})^*"] & X_{\lambda(t)t^{-s}}^G, \end{tikzcd} \tag{5.3}$$

and

$$\begin{tikzcd} r_\alpha(T^*T) : X_{t^{p-s}}^G \arrow[r, "J^g_{p-s}"] & X_g^H \arrow[r, "r_\alpha(T^*T)"] & X_l^H \arrow[r, "(J^{1/l}_{1/\lambda(t)t^{-s}})^*"] & X_{\lambda(t)t^{-s}}^G \end{tikzcd} \tag{5.4}$$

Note that norm bounds for the embeddings J_{p-s}^g were given in Proposition 4.1. Also we established results for bounding the noise term in Section 4.4 and tools for determining the qualification of regularizations in Section 4.1, such that below we may use previous results as a tool box.

It is remarkable to note that in all cases considered below, the parameter choice which yields the order optimal error bound is independent of the underlying norm and is obtained by

$$\sqrt{\frac{\alpha_*}{m^2}}\, g(\frac{\alpha_*}{m^2}) = \frac{\delta}{mE}, \quad 0 < \delta \le mEb^{p-s}\Psi_s(b). \tag{5.5}$$

5.1. **Bounding the error in $\|\cdot\|_s,\ s \le p$.** The error analysis is particularly simple in case $\lambda(t) := t^s$, where we have to restrict ourselves to $s \le p$, to make sure that $x^\dagger \in X_{t^s}^G$. In this case the analysis is similar to the analysis for ordinary Tikhonov regularization with $s = 0$. Indeed, if $\lambda(t) := t^s$, then $l(t) \equiv 1$ and the corresponding diagram from (5.3) and (5.4) reduces to

$$\begin{tikzcd} g_\alpha(T^*T) : X_{\sqrt{t}}^H \arrow[rr, "g_\alpha(T^*T)"] & & X \\ r_\alpha(T^*T) : X_{t^{p-s}}^G \arrow[r, "J^g_{p-s}"] & X_g^H \arrow[r, "r_\alpha(T^*T)"] & X. \end{tikzcd}$$

We state the following result.

Proposition 5.1. *Suppose $s \le p$. Under assumptions A.1, A.4.1 and A.5 we have*

$$\|x_\alpha^\delta - x^\dagger\|_s \le \gamma E g(\alpha/m^2) + \gamma_* \frac{\delta/m}{\sqrt{\alpha/m^2}}. \tag{5.6}$$

If we let $\alpha_* = \alpha_*(\delta)$ a priori *be chosen according to (5.5) then*

$$\|x_\alpha^\delta - x^\dagger\|_s \le (\gamma + \gamma_*)\, E g(\delta/mE). \tag{5.7}$$

Proof. The noise term $\|g_\alpha(T^*T)\colon X^H_{\sqrt{t}} \to X\|$ was bounded in (4.12) and we need to bound $\|r_\alpha(T^*T)\colon X^H_g \to X\|$. Under assumptions A.4.1 and A.5 we deduce from Corollary 4.1 that

$$\|r_\alpha(T^*T)\colon X^H_g \to X\| = \sup_{0<t\le a} |r_\alpha(t)|\, g(t/m^2) \le \gamma g(\alpha/m^2).$$

The remaining assertion (5.7) follows from that. □

5.2. **Bounding the error in $\|\cdot\|_{1/\varrho}$.** As could be seen from the application of the interpolation inequality in the proof of Corollary 3.1 as limiting cases we used bounds in the norm $\|\cdot\|_{1/\varrho}$ and $\|\cdot\|_p$, the former is the weakest norm to analyze the error. Here $\lambda(t)t^{-s} = 1/\Psi_s(t)$ and by duality, assumption (4.9) implies that we can let $l(t) := t^{-1/2}$ in the diagrams (5.3) and (5.4), which now specifies to

$$\begin{aligned}
g_\alpha(T^*T) : X^H_{\sqrt{t}} &\xrightarrow{g_\alpha(T^*T)} X^H_{1/\sqrt{t}} \xrightarrow{(J^{1/2}_{\Psi_s(t)})^*} X^G_{1/\Psi_s(t)},\\
r_\alpha(T^*T) : X^G_{t^{p-s}} &\xrightarrow{J^g_{p-s}} X^H_g \xrightarrow{r_\alpha(T^*T)} X^H_{1/\sqrt{t}} \xrightarrow{(J^{1/2}_{\Psi_s(t)})^*} X^G_{1/\Psi_s(t)}.
\end{aligned}$$

The embeddings can be bounded under Assumption A.5 and the corresponding A.1 and A.2. The noise amplification was bounded in Lemma 4.3 and we are left with bounding the pure residual.

Lemma 5.1.

$\mathbf{p \le s}$: *Under Assumption A.4.1 we have*

$$\|r_\alpha(T^*T)\colon X^H_g \to X^H_{1/\sqrt{t}}\| \le \gamma\sqrt{\alpha/m^2}\, g(\alpha/m^2). \tag{5.8}$$

$\mathbf{s < p}$: *If A.5 and the stronger Assumption A.4.2 hold true then*

$$\|r_\alpha(T^*T)\colon X^H_g \to X^H_{1/\sqrt{t}}\| \le \gamma\sqrt{\alpha/m^2}\, g(\alpha/m^2). \tag{5.9}$$

Proof. Plainly

$$\|r_\alpha(T^*T)\colon X^H_g \to X^H_{1/\sqrt{t}}\| = \sup_{0<t\le a} |r_\alpha(t)|\, g(t/m^2)\sqrt{t/m^2}.$$

In the first case the function $g(t)\sqrt{t}/\sqrt{t} = g(t)$ is decreasing. Since by Lemma 4.2 the function $g(t)\sqrt{t}$ is a proper index function we can apply Lemma 4.1 with $D = 1$ to complete the proof.

In the second case, the function g^2 obeys Assumption A.3, thus the second assertion of Corollary 4.1 applies (with constant $D_{g^2} = 1$) and the proof is completed in this case. □

We can summarize the above bounds to state the final error estimate.

Proposition 5.2. *Let assumptions A.1 and A.5 hold true.*

$\mathbf{p} \le \mathbf{s}$: *Under A.2 and if the regularization g_α obeys A.4.1 then*

$$\|x_\alpha^\delta - x^\dagger\|_{1/\varrho} \le \gamma \frac{M}{m} E \sqrt{\alpha/m^2}\, g(\alpha/m^2) + \beta \frac{\delta}{m}.$$

$\mathbf{s} \le \mathbf{p}$: *If the regularization g_α obeys A.4.2 then*

$$\|x_\alpha^\delta - x^\dagger\|_{1/\varrho} \le \gamma E \sqrt{\alpha/m^2}\, g(\alpha/m^2) + \beta \frac{\delta}{m}.$$

If we let $\alpha_* = \alpha_*(\delta)$ a priori *be chosen as in (5.5) then we obtain*

$\mathbf{p} \le \mathbf{s}$: $\|x_{\alpha_*}^\delta - x^\dagger\|_{1/\varrho} \le (\beta + \frac{M}{m}\gamma)\delta/m.$
$\mathbf{s} \le \mathbf{p}$: $\|x_{\alpha_*}^\delta - x^\dagger\|_{1/\varrho} \le (\beta + \gamma)\delta/m.$

Thus in this weak norm the rate of recovering the solution behaves as if the problem were well-posed, because one cannot beat the intrinsic inaccuracy of level δ in the data y^δ.

5.3. **Bounding the error in $\|\cdot\|_p$.** The other benchmark case is met for $\lambda(t) := t^p$, because this is the maximal norm at which we may analyze the error. In this case we shall use the factorizations (5.3) and (5.4) with the function $l(t) := g(t)$, i.e.,

$$\begin{aligned} &g_\alpha(T^*T) : X^H_{\sqrt{t}} \xrightarrow{g_\alpha(T^*T)} X^H_g \xrightarrow{(J^{1/g}_{s-p})^*} X^G_{t^{p-s}}, \\ &r_\alpha(T^*T) : X^G_{t^{p-s}} \xrightarrow{J^g_{p-s}} X^H_g \xrightarrow{r_\alpha(T^*T)} X^H_g \xrightarrow{(J^{1/g}_{s-p})^*} X^G_{t^{p-s}}. \end{aligned} \tag{5.10}$$

Again, under Assumption A.5, in conjunction with the appropriate assumptions A.1 and A.2 we can bound the norms of the embeddings. The noise part was bounded in Proposition 4.2. The residual bound is particularly simple, because

$$\|r_\alpha(T^*T)\colon X^H_g \to X^H_g\| = \sup_{0<t\le a} |r_\alpha(t)|\,(1/g)(t/m^2) g(t/m^2) \le \gamma,$$

and we summarize the above analysis as follows.

Proposition 5.3. *Under assumptions A.1, A.2, A.5 and A.6 the following bound is valid in both cases $0 < p \le s$ and $0 \le s < p$:*

$$\|x_\alpha^\delta - x^\dagger\|_p \le \gamma \frac{M}{m} E + \max\{\beta, \gamma^*\} \frac{\delta/m}{\sqrt{\alpha/m^2}\, g(\alpha/m^2)}.$$

Thus in either case, if α_* is chosen according to (5.5), then

$$\|x_{\alpha_*}^\delta - x^\dagger\|_p \le \left(\gamma \frac{M}{m} + \max\{\beta, \gamma^*\}\right) E. \tag{5.11}$$

Here it is interesting to note that no rate of reconstruction can be deduced. Heuristically this is clear, because in this strong norm the effective solution smoothness, as this was introduced in [16] and generalized to variable Hilbert scales in [14], is equal to zero.

5.4. **Bounding the error in $\| \cdot \|_X$.** Finally we turn to estimating the error in the original Hilbert space X. This could be done similar to the analysis above. However we feel it interesting to use the interpolation inequality as presented in Theorem 3.1. We state the main result of our error analysis, and we use the function $\Psi_p(t)$ from (2.5).

Theorem 5.1. *Let the assumptions A.1, A.2, A.5 and A.6 hold true. Assume furthermore that* $t \mapsto \left[\left(\Psi_p^2\right)^{-1}(t) \right]^{2p}$ *is concave. Let* $\alpha_* = \alpha_*(\delta)$ a priori *be chosen as in (5.5). If*

$\mathbf{p} \leq \mathbf{s}$: *either g_α has qualification as in A.4.1,*
$\mathbf{s} \leq \mathbf{p}$: *or g_α has qualification as in A.4.2,*

then

$$\sup_{x^\dagger \in G_{p,E}} \|x^\delta_{\alpha_*} - x^\dagger\| \leq \left(\gamma \frac{M}{m} + \max\{\beta, \gamma^*\} \right) E \left[\Psi_p^{-1}\left(\frac{\delta}{mE} \right) \right]^p .$$

In both cases the order of the bound from (2.6) is obtained.

Proof. We shall apply the interpolation inequality (3.1) with parameters $\theta(t) = \varphi(t) = t^p$ and $\psi(t) = \Psi_p(t)$ to the element $v := x^\delta_{\alpha_*} - x^\dagger$. Under the concavity assumption this implies

$$\|x^\delta_{\alpha_*} - x^\dagger\| \leq \|x^\delta_{\alpha_*} - x^\dagger\|_p \left[\Psi_p^{-1}\left(\frac{\|x^\delta_{\alpha_*} - x^\dagger\|_{1/\varrho}}{\|x^\delta_{\alpha_*} - x^\dagger\|_p} \right) \right]^p . \tag{5.12}$$

In the high order case, let us temporarily denote by $C_p := \gamma \frac{M}{m} + \max\{\beta, \gamma^*\}$ and $C_{1/\varrho} := (\beta + \gamma \frac{M}{m})$ the constants in the bounds from Propositions 5.2 and 5.3, respectively. Because $C_{1/\varrho} \leq C_p$ we use monotonicity to deduce from (5.12) that

$$\|x^\delta_{\alpha_*} - x^\dagger\| \leq C_p E \left[\Psi_p^{-1}\left(\frac{\delta}{mE} \right) \right]^p . \tag{5.13}$$

In the low order case, we have to replace $C_{1/\varrho}$ by the corresponding $C_{1/\varrho} := \beta + \gamma$ and we obtain the corresponding bound. □

Remark 9. It is worth-wile to note, that in the high order case $0 < p \leq s$ we could have applied the interpolation inequality (3.1) with $\theta(t) = t^s$ and then use the bounds from Proposition 5.1. In this case the bounds from Theorem 5.1 can be obtained, with different constants, without using Assumption A.2.

In all cases covered by Theorem 5.1 the validity of Assumption A.5 is crucial, thus either $s > p$ and the function $1/g^2$ is operator monotone or $s \leq p$ and the function g^2 is assumed to be operator monotone. In the context of Natterer's result as mentioned in (1.4) and in the light of the discussion in Example 4.5 (finitely smoothing case), the latter corresponds to the restriction $s \geq (p - a)/2$ as supposed to hold for (1.5).

If regularization is carried out with $s = p$ then $g(t) \equiv 1$ and the theorem holds without additional assumptions on operator monotonicity.

Remark 10. Let us discuss the Landweber iteration as introduced after Definition 3 for the finitely smoothing case of Example 4.5 in some detail. Since the number n of iterations and α are related by $\alpha = 1/n$ we obtain from the parameter choice (5.5) that for

$$n = \mathcal{O}\Big(\delta^{-\frac{2(a+s)}{a+p}}\Big)$$

order optimal error bounds can be guaranteed. We see that for smaller s-values the number n of necessary iterations decreases. From the viewpoint of complexity the aim consists therefore in working with s-values as small as possible. The constraints that g^2 is operator monotone and that $g(t)/\sqrt{t}$ is decreasing imply that $s_0 := (p-a)/2$ is the smallest possible s-value leading to order optimal error bounds. This s-value may be even negative in case $p < a$ and requires only $n_0 = O(\delta^{-1})$ iterations. For order optimal convergence rate results in case of negative s-values, that require even fewer iterations see [6].

6. A POSTERIORI CHOICE OF THE REGULARIZATION PARAMETER

If the constants m and E in the *a priori* parameter choice (5.5) are unknown, then the parameter choice

$$\alpha = \left(\frac{\delta}{c_2}\right)^2 \left[\Psi_p^{-1}\left(\frac{\delta}{c_1 c_2}\right)\right]^{2(s-p)} \tag{6.1}$$

may be used where c_1 and c_2 are positive constants guessing m and E, respectively. For the parameter choice (6.1), the results of the Theorem 5.1 still hold true where the respective error bounds have to be replaced by the order optimal error bound

$$\|x_\alpha^\delta - x^\dagger\| \le cE\left[\Psi_p^{-1}\left(\frac{\delta}{mE}\right)\right]^p \quad \text{with some} \quad c \ge 1\,. \tag{6.2}$$

If not only m and E, but also p and ψ_p are unknown, then *a posteriori* rules for choosing the regularization parameter α have to be used.

6.1. **Using the discrepancy principle.** In Morozov's discrepancy principle (see [21]) the regularization parameter $\alpha = \alpha_D$ is chosen as the solution of the nonlinear equation

$$d(\alpha) := \|Ax_\alpha^\delta - y^\delta\| = C\delta \tag{6.3}$$

with some constant $C > 1$. Actually we shall need C to exceed the constant γ from (4.2). We will assume without further mentioning that the nonlinear equation (6.3) possesses a unique solution. Conditions guaranteeing this may be found, e.g., in [32]. We shall show that for $\alpha = \alpha_D$ the order optimal error bounds from Theorem 5.1 still hold true under analogous assumptions. This result will be obtained similarly to the previous analysis by using the interpolation inequality (3.1), after bounding the errors on $\|\cdot\|_{1/\varrho}$ and $\|\cdot\|_p$, respectively. In a first auxiliary proposition we provide some lower bound for the regularization parameter $\alpha = \alpha_D$.

Proposition 6.1. *Let x_α^δ be the regularized approximation from (4.6) and let $\alpha = \alpha_D$ be chosen by the discrepancy principle (6.3). There is a constant $\tilde{k} > 0$ such that under A.5 and the assumptions of Lemma 5.1 we have*

$$(C-\gamma)\delta \le \tilde{k}E\sqrt{\alpha_D}\,g(\alpha_D/m^2). \tag{6.4}$$

Proof. We first note that by (4.6) the discrepancy can be rewritten as

$$Ax^\delta_{\alpha_D} - y^\delta = AG^s g_{\alpha_D}(T^*T)T^*y^\delta - y^\delta = r_{\alpha_D}(TT^*)y^\delta.$$

Therefore we obtain for $y := Ax^\dagger$ that

$$C\delta = \|Ax^\delta_{\alpha_D} - y^\delta\| = \|r_{\alpha_D}(TT^*)y^\delta\| \le \|r_{\alpha_D}(TT^*)(y - y^\delta\| + \|r_{\alpha_D}(TT^*)y\|,$$

hence $(C-\gamma)\delta \le \|r_{\alpha_D}(TT^*)y\|$, and we need to upper bound $\|r_{\alpha_D}(TT^*)y\|$. As indicated, we shall make use of Lemma 5.1 as follows. Let $x^\dagger = G^p v$.

$$\begin{aligned}\|r_{\alpha_D}(TT^*)y\| &= \|r_{\alpha_D}(TT^*)TG^{p-s}v\| \\ &\le E\|J^g_{p-s}\colon X^G_{p-s} \to X^H_g\|\|r_{\alpha_D}(TT^*)T\colon X^H_g \to X\| \\ &= E\|J^g_{p-s}\colon X^G_{p-s} \to X^H_g\|\|r_{\alpha_D}(T^*T)\colon X^H_g \to X^H_{1/\sqrt{t}}\|.\end{aligned}$$

The norm of the embedding can be bounded under A.5 and the second norm was bounded in Lemma 5.1, which completes the proof of the proposition. □

Corollary 6.1. *Let x^δ_α be the regularized approximation from (4.6) and let $\alpha = \alpha_D$ be chosen by the discrepancy principle (6.3). There is a constant $k > 0$ such that under the assumptions of Proposition 5.3 we have*

$$\|x^\delta_{\alpha_D} - x^\dagger\|_p \le kE. \tag{6.5}$$

Proof. The required bound is obtained, if we insert the lower bound from (6.4) into the estimate from Proposition 5.3. □

We now establish an appropriate bound in $\|\cdot\|_{1/\varrho}$.

Proposition 6.2. *Let $\alpha = \alpha_D$ be chosen by the discrepancy principle (6.3). Under Assumption A.1 we have*

$$\|x^\delta_{\alpha_D} - x^\dagger\|_{1/\varrho} \le \frac{C+1}{m}\delta. \tag{6.6}$$

Proof. This is immediate from

$$\begin{aligned}\|x^\delta_{\alpha_D} - x^\dagger\|_{1/\varrho} = \|\rho(G)(x^\delta_{\alpha_D} - x^\dagger)\| &\le \frac{1}{m}\|Ax^\delta_{\alpha_D} - Ax^\dagger\| \\ &\le \frac{1}{m}\left(\|Ax^\delta_{\alpha_D} - y^\delta\| + \|y - y^\delta\|\right) \le \frac{(C+1)\delta}{m}.\end{aligned}$$

□

We can now use the interpolation inequality (3.1) in the same way as in the proof of Theorem 5.1 to conclude the main error bound under the discrepancy principle.

Theorem 6.1. *Let x^δ_α be defined by the general regularization scheme (4.6) and let $\alpha = \alpha_D$ be chosen by the discrepancy principle (6.3). Assume the solution smoothness obeys $\|x^\dagger\|_p \le E$ and that $t \mapsto \left[\left(\Psi_p^2\right)^{-1}(t)\right]^{2p}$ is concave. Under assumptions A.1–A.6 there is a constant $K < \infty$ such that*

$$\|x^\delta_{\alpha_D} - x^\dagger\| \le KE\left[\Psi_p^{-1}\left(\frac{\delta}{mE}\right)\right]^p. \tag{6.7}$$

6.2. **Using the Lepskiĭ principle.** Here we recall the a posteriori choice according to the Lepskiĭ, or balancing principle, as e.g. outlined in [19, Appendix]. For a recent account on this principle we refer to [15]. It is based on a valid bound on the noise term in dependence of α. Typically such bound would depend on properties of the link condition A.1, however the constants m and the link function ϱ need not be known to us and our goal is to use the Lepskiĭ principle without such knowledge. It turns out that this can be done in the $\|x\|_A := \|Ax\|,\ x \in X$, as can be seen from estimate (6.8), below. Note, that by definition of the operator T^*T it holds true that $\|x\|_A = m\|x\|_{X^H_{\sqrt{t}}}$, and the same arguments used for proving Proposition 5.2 apply. For the convenience of the reader we shall repeat the basic arguments.

First, using (4.7) and the definition of β from (4.3) we bound the noise term as

$$\|A(x_\alpha^\delta - x_\alpha)\| = \|Tg_\alpha(T^*T)T^*(y - y^\delta)\| \le \beta\delta, \tag{6.8}$$

which provides us with a valid noise bound as $\beta\delta$. It remains to bound the bias.

Lemma 6.1. *Suppose $x^\dagger \in G_p(E)$ and that assumptions A.1 and A.5 hold true.*

$\mathbf{p \le s}$: *Under A.2 and if the regularization g_α obeys A.4.1 then*

$$\|A(x_\alpha - x^\dagger)\| \le \gamma\frac{M}{m}E\sqrt{\alpha}g(\alpha/m^2).$$

$\mathbf{s \le p}$: *If the regularization g_α obeys A.4.2 then*

$$\|A(x_\alpha - x^\dagger)\| \le \gamma E\sqrt{\alpha}g(\alpha/m^2).$$

Sketch of the proof. Using representation (4.8) we can bound

$$\begin{aligned}\|A(x_\alpha - x^\dagger)\| &\le E\|Tg_\alpha(T^*T)G^{p-s}\| \\ &\le E\|Tg_\alpha(T^*T)\colon X_g^H \to Y\|\|J^g_{p-s}\colon X^G_{t^{p-s}} \to X^H_g\|.\end{aligned}$$

Now we can continue using the same arguments as in the proof of Lemma 5.1. □

The choice of regularization parameter according to Lepskiĭ is as follows. Starting from some small enough α_1, typically $\alpha_1 \asymp \delta^2$, and fixing some parameter $q > 1$ we let

$$M := \{\alpha_j, \qquad j = 1, 2, \dots, N\},$$

where $\alpha_j := \alpha_1 q^{j-1}$, and N is determined as smallest integer such that $\alpha_N > 1$. The Lepskiĭ principle chooses the parameter, say α_L, from this finite set M. The Lepskiĭ index j_L is determined as

$$j_L := \max\left\{j, \quad \|A(x_{\alpha_j^\delta} - x_{\alpha_i^\delta})\| \le 4\beta\delta \text{ for all } i < j\right\}.$$

We denote the corresponding regularization parameter

$$\alpha_L := \alpha_{j_L}. \tag{6.9}$$

The following remarkable properties of this choice of parameter are known from [19, Appendix]. To this end we recall, and specify to the present context, that an index function $\Phi(\alpha)$ is called admissible for $x^\dagger$ if there is a valid error bound as

$$\|A(x_\alpha^\delta - x^\dagger)\| \le \Phi(\alpha) + \beta\delta, \quad \alpha \in M, \tag{6.10}$$

and $\Phi(\alpha_1) \leq \beta\delta$, where the latter condition can be ensured by using α_1 small enough. Note that by Lemma 6.1 the function $\alpha \to \gamma EM/m\sqrt{\alpha}g(\alpha/m^2)$ is admissible under the corresponding assumptions.

For discussing properties of the Lepskiĭ principle for choosing α we introduce some auxiliary index

$$j_* := \max\{j, \quad \text{there is an admiss. } \Phi \text{ such that } \Phi(\alpha_j) \leq \beta\delta\}. \tag{6.11}$$

With these preparations we can state the following two properties of the choice α_L as established in [19, Appendix, Lemma 4], precisely,

$$j_L \geq j_* \tag{6.12}$$

and

$$\|A(x^\delta_{\alpha_L} - x^\dagger)\| \leq 6\beta\delta. \tag{6.13}$$

We shall use the first property (6.12) to draw a conclusion similar to Proposition 6.1.

Proposition 6.3. *Assume that $x^\dagger \in G_p(E)$. Let x^δ_α be the regularized approximation from (4.6) and let $\alpha = \alpha_L$ be chosen by the Lepskiĭ principle (6.9). Under A.1– A.6 there is a constant $\tilde{k} > 0$ such that*

$$\delta \leq \tilde{k}E\sqrt{\alpha_L/m^2}g(\alpha_L/m^2). \tag{6.14}$$

Proof. Since $q\alpha_L > \alpha_L$ property (6.12) implies that $\beta\delta \leq \Phi(q\alpha_L)$ for every admissible function. Therefore Lemma 6.1 implies that

$$\beta\delta \leq \gamma M/mE\sqrt{q\alpha_L}g(q\alpha_L/m^2).$$

Now we distinguish the high order and low order cases. In the high order case the function g is decreasing and we obtain

$$\beta\delta \leq \sqrt{q}\gamma M/mE\sqrt{\alpha_L}g(\alpha_L/m^2).$$

In the low order case the function $t \mapsto g(t)/\sqrt{t}$ was decreasing and we obtain

$$\beta\delta \leq \gamma M/mE(q\alpha_L)g(q\alpha_L/m^2)/\sqrt{q\alpha_L} \leq q\gamma M/mE\sqrt{\alpha_L}g(\alpha_L/m^2),$$

and the proof is complete with $\tilde{k} = q\gamma M/(m\beta)$. □

As in Section 6.1 we conclude

Corollary 6.2. *Let x^δ_α be the regularized approximation from (4.6) and let $\alpha = \alpha_L$ be chosen by the Lepskiĭ principle (6.9). If $x^\dagger \in G_p(E)$ and assumptions A.1– A.6 hold true then there is a constant $k > 0$ such that*

$$\|x^\delta_{\alpha_L} - x^\dagger\|_p \leq kE. \tag{6.15}$$

Together with the noise bound given in (6.13) we can finally establish the order optimality of the Lepskiĭ principle.

Theorem 6.2. *Let x_α^δ be defined by the general regularization scheme (4.6) and let $\alpha = \alpha_L$ be chosen by the Lepskiĭ principle (6.9). Assume the solution smoothness obeys $\|x^\dagger\|_p \le E$ and that $t \mapsto \left[\left(\Psi_p^2\right)^{-1}(t)\right]^{2p}$ is concave. Under assumptions A.1–A.6 there is a constant $K < \infty$ such that*

$$\|x_{\alpha_L}^\delta - x^\dagger\| \le KE\left[\Psi_p^{-1}\left(\frac{\delta}{mE}\right)\right]^p. \tag{6.16}$$

7. Discussion of the assumptions

As can be seen from the proofs, the assumptions A.1–A.6 have different impact at different places and we find it worth-wile to discuss their relevance.

Clearly, assumptions A.1 and A.2 relate the smoothness of the underlying true solution $x^\dagger$ to the scale generated by $H := A^*A/m$. Without such *linking conditions* the regularization theory is not applicable, and no rates can be obtained.

The Assumptions A.4 ensure a *minimal qualification* of the initial regularization as given through the families $g_\alpha(t),\ 0 < \alpha < \infty$. This requirement is very weak and it is provided by most regularization schemes.

The crucial *geometric assumption* is introduced by A.5, which in turn uses A.3. The use of such geometric assumptions is the main goal of the present analysis.

Assumption A.6 is stronger than the corresponding normalization (4.1) in Definition 2. As explained in § 4.4 this is necessary to *bound the noise amplification* in the strong norm of $X_{\sqrt{t}}^H$. If we would have available an upper bound for the smoothness of $x^\dagger$ then this might be relaxed. For instance, if we knew some p_0 for which the smoothness $p \le p_0$ of $x^\dagger$ is bounded from above, and if we would restrict ourselves to high-order regularization with $s \ge p_0$, then A.6 can be omitted.

References

[1] Jöran Bergh and Jörgen Löfström. *Interpolation spaces. An introduction.* Springer-Verlag, Berlin, 1976. Grundlehren der Mathematischen Wissenschaften, No. 223.

[2] Rajendra Bhatia. *Matrix analysis.* Springer-Verlag, New York, 1997.

[3] Albrecht Böttcher, Bernd Hofmann, Ulrich Tautenhahn, and Masahiro Yamamoto. Convergence rates for Tikhonov regularization from different kinds of smoothness conditions. *Applicable Analysis*, 85, 2006. published electronically as DOI: 10.1080/00036810500474838.

[4] William F. Donoghue, Jr. The interpolation of quadratic norms. *Acta Math.*, 118:251–270, 1967.

[5] William F. Donoghue, Jr. *Monotone matrix functions and analytic continuation.* Springer-Verlag, New York, 1974.

[6] Herbert Egger and Andreas Neubauer. Preconditioning Landweber iteration in Hilbert scales. *Numer. Math.*, 101(4):643–662, 2005.

[7] Heinz W. Engl, Martin Hanke, and Andreas Neubauer. *Regularization of inverse problems.* Kluwer Academic Publishers Group, Dordrecht, 1996.

[8] Frank Hansen and Gert Kjaergård Pedersen. Jensen's inequality for operators and Löwner's theorem. *Math. Ann.*, 258(3):229–241, 1981/82.

[9] Markus Hegland. An optimal order regularization method which does not use additional smoothness assumptions. *SIAM J. Numer. Anal.*, 29(5):1446–1461, 1992.

[10] Markus Hegland. Variable Hilbert scales and their interpolation inequalities with applications to Tikhonov regularization. *Appl. Anal.*, 59(1-4):207–223, 1995.

[11] Bernd Hofmann. *Regularization for applied inverse and ill-posed problems*, volume 85 of *Teubner-Texte zur Mathematik [Teubner Texts in Mathematics]*. BSB B. G. Teubner Verlagsgesellschaft, Leipzig, 1986.

[12] Fengshan Liu and M. Zuhair Nashed. Tikhonov regularization of nonlinear ill-posed problems with closed operators in Hilbert scales. *J. Inverse Ill-Posed Probl.*, 5(4):363–376, 1997.

[13] Alfred Karl Louis. *Inverse und schlecht gestellte Probleme*. B. G. Teubner, Stuttgart, 1989.

[14] Peter Mathé. Degree of ill-posedness of statistical inverse problems. Preprint 954, WIAS, August 2004.

[15] Peter Mathé. The Lepskiĭ principle revisited. *Inverse Problems*, 22, 2006. publ. electronically as DOI:10.1088/0266-5611/22/3/L02.

[16] Peter Mathé and Sergei V. Pereverzev. Optimal discretization of inverse problems in Hilbert scales. Regularization and self-regularization of projection methods. *SIAM J. Numer. Anal.*, 38(6):1999–2021, 2001.

[17] Peter Mathé and Sergei V. Pereverzev. Discretization strategy for linear ill-posed problems in variable Hilbert scales. *Inverse Problems*, 19(6):1263–1277, 2003.

[18] Peter Mathé and Sergei V. Pereverzev. Geometry of linear ill-posed problems in variable Hilbert scales. *Inverse Problems*, 19(3):789–803, 2003.

[19] Peter Mathé and Sergei V. Pereverzev. Regularization of some linear ill–posed problems with discretized random noisy data. to appear Math. Comput., 2005.

[20] Peter Mathé and Ulrich Tautenhahn. Error bounds for regularization methods in Hilbert scales by using operator monotonicity. *Far East J. Math. Sci.*, 2006. to appear.

[21] Vitali A. Morozov. On the solution of functional equations by the method of regularization. *Soviet Math. Dokl.*, 7:414–417, 1966.

[22] M. Thamban Nair. On Morozov's method for Tikhonov regularization as an optimal order yielding algorithm. *Z. Anal. Anwendungen*, 18(1):37–46, 1999.

[23] M. Thamban Nair. Optimal order results for a class of regularization methods using unbounded operators. *Integral Equations Operator Theory*, 44(1):79–92, 2002.

[24] M. Thamban Nair, Sergei V. Pereverzev, and Ulrich Tautenhahn. Regularization in Hilbert scales under general smoothing conditions. *Inverse Problems*, 21:1851–1869, 2005.

[25] Frank Natterer. Error bounds for Tikhonov regularization in Hilbert scales. *Applicable Anal.*, 18(1-2):29–37, 1984.

[26] Andreas Neubauer. An a posteriori parameter choice for Tikhonov regularization in Hilbert scales leading to optimal convergence rates. *SIAM J. Numer. Anal.*, 25(6):1313–1326, 1988.

[27] Andreas Neubauer. Tikhonov regularization of nonlinear ill-posed problems in Hilbert scales. *Appl. Anal.*, 46(1-2):59–72, 1992.

[28] Ulrich Tautenhahn. Error estimates for regularization methods in Hilbert scales. *SIAM J. Numer. Anal.*, 33(6):2120–2130, 1996.

[29] Ulrich Tautenhahn. On a general regularization scheme for nonlinear ill-posed problems. II. Regularization in Hilbert scales. *Inverse Problems*, 14(6):1607–1616, 1998.

[30] Andrey N. Tikhonov and Vasiliy Y. Arsenin. *Solutions of ill-posed problems*. V. H. Winston & Sons, Washington, D.C.: John Wiley & Sons, New York, 1977. Translated from the Russian, Preface by translation editor Fritz John, Scripta Series in Mathematics.

[31] Mitsuru Uchiyama. A new majorization between functions, polynomials, and operator inequalities. *J. Funct. Anal.*, 231(1):221–244, 2006.

[32] Gennadi M. Vaĭnikko and Alexander Yu. Veretennikov. *Iteratsionnye protsedury v nekorrektnykh zadachakh*. "Nauka", Moscow, 1986.

WEIERSTRASS INSTITUTE FOR APPLIED ANALYSIS AND STOCHASTICS, MOHRENSTRASSE 39,
10117 BERLIN, GERMANY

E-mail address: `mathe@wias-berlin.de`

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF APPLIED SCIENCES ZITTAU/GÖRLITZ, P.O.BOX
1454, 02754 ZITTAU, GERMANY

E-mail address: `u.tautenhahn@hs-zigr.de`