INSTITUTE OF PHYSICS PUBLISHING INVERSE PROBLEMS

Inverse Problems **21** (2005) 1851–1869 doi:10.1088/0266-5611/21/6/003

Regularization in Hilbert scales under general smoothing conditions

M T Nair[1], **S V Pereverzev**[2] **and U Tautenhahn**[3]

[1] Department of Mathematics, Indian Institute of Technology Madras, Chennai 600 036, India
[2] Johann Radon Institute for Computational and Applied Mathematics, Austrian Academy of Sciences, Altenbergstrasse 69, 4040 Linz, Austria
[3] Department of Mathematics, University of Applied Sciences Zittau/Görlitz, PO Box 1455, 02755 Zittau, Germany

E-mail: mtnair@iitm.ac.in, sergei.pereverzyev@oeaw.ac.at and u.tautenhahn@hs-zigr.de

Received 5 July 2005, in final form 8 September 2005
Published 3 October 2005
Online at stacks.iop.org/IP/21/1851

Abstract

For solving linear ill-posed problems regularization methods are required when the available data include some noise. In the present paper regularized approximations are obtained by a general regularization scheme in Hilbert scales which include well-known regularization methods such as the method of Tikhonov regularization and its higher-order forms, spectral methods, asymptotical regularization and iterative regularization methods. For both the cases of high- and low-order regularization, we study *a priori* and *a posteriori* rules for choosing the regularization parameter and provide order optimal error bounds that characterize the accuracy of the regularized approximations. These error bounds have been obtained under general smoothing conditions. The results extend earlier results and cover the case of finitely and infinitely smoothing operators. The theory is illustrated by a special ill-posed deconvolution problem arising in geoscience.

1. Introduction

Ill-posed problems arise in several contexts and have important applications in science and engineering (see, e.g., [3, 4, 8, 22]). In this paper we consider ill-posed problems

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

where $A : X \to Y$ is a bounded linear operator between infinite dimensional real Hilbert spaces X and Y with non-closed range $\mathcal{R}(A)$. We shall denote the inner product and the corresponding norm on the Hilbert spaces by $\langle\cdot,\cdot\rangle$ and $\|\cdot\|$ respectively. We assume throughout the paper that the operator A is injective and that y belongs to $\mathcal{R}(A)$ so that (1.1) has a unique solution $x^\dagger \in X$. Suppose that the available data are $y^\delta \in Y$ in place of the exact data y such that

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

for some known noise level δ. Since $\mathcal{R}(A)$ is assumed to be non-closed, the solution $x^\dagger$ does not depend continuously on the data. Hence, the numerical treatment of problems (1.1) and (1.2) requires the application of special regularization methods. In the method of Tikhonov regularization in Hilbert scales a regularized approximation x_α^δ is defined as the solution of the minimization problem

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

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

In many practical problems the operator B which influences the properties of the regularized approximation is chosen to be a differential operator in some appropriate function spaces, e.g., L^2-spaces. In [16] Natterer has shown that under the assumptions

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

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

$$\left\|x_\alpha^\delta - x^\dagger\right\| = O(\delta^{p/(a+p)}) \qquad \text{for} \quad s \geqslant (p-a)/2 \tag{1.5}$$

in the case that α is chosen *a priori* by $\alpha = c\delta^{2(a+s)/(a+p)}$ with some constant $c > 0$.

In the meantime regularization in Hilbert scales became quite popular; see, e.g., [13, 17] where method (1.3) has been studied with α chosen from Morozov's discrepancy principle, [14, 20] where method (1.3) has been generalized to a general regularization scheme, [9, 14] where extensions to the case of infinitely smoothing operators A have been treated or [3, 7, 18, 21] in which extensions to the nonlinear case may be found. The main aim of this paper is to derive results on order optimal convergence rates in cases of proper *a priori* and *a posteriori* parameter choice strategies. Our analysis has been done

(i) for regularization methods that are more general than (1.3) and
(ii) in the case of smoothing conditions that are more general than (1.4).

The paper is organized as follows. In section 2 we introduce the smoothing conditions; see assumptions A1 and A2 that characterize the smoothness of the unknown solution $x^\dagger$ of problem (1.1) and the smoothing properties of the operator A relative to the operator B^{-1} which can be quite independent of A. Under these conditions and (1.2), some estimate is provided that characterizes the best possible worst case error for identifying $x^\dagger$ from noisy data y^δ. In section 3 we introduce a general regularization scheme in Hilbert scales and consider the case of known solution smoothness. We prove that by standard regularization methods with proper stabilization in dependence of the known solution smoothness, order optimal error bounds can be guaranteed provided the regularization parameter has been chosen properly, either *a priori* or *a posteriori* using Morozov's discrepancy principle. In sections 4 and 5 we study the case of unknown solution smoothness. In this case we divide our study into two subcases. In the first subcase we consider high-order regularization in which much smoothness is introduced into the regularization procedure, and the second subcase is concerned with low-order regularization in which little smoothness is introduced. In both subcases order optimal error bounds can be guaranteed provided the regularization parameter has been chosen properly. From the viewpoint of complexity, the subcase of low-order regularization seems to be especially important. In section 6 we discuss a possible application of our results to the deconvolution problem arising in geoscience in the context of models with a non-Wiener filter design. Final remarks are presented in section 7.

2. Optimality and order optimality

Since problem (1.1) is ill-posed, for the stable reconstruction of the solution $x^\dagger$ of problem (1.1) with inexact data y^δ satisfying (1.2) additional information is required. In this regard, we assume in this paper that we have the following pieces of information:

(i) information concerning the smoothness of $x^\dagger$ and
(ii) information concerning the smoothing property of the operator A.

We formulate our additional information in terms of some densely defined unbounded self-adjoint strictly positive operator B with its domain and range in X. We introduce a *Hilbert scale* $(X_r)_{r\in\mathbf{R}}$ induced by the operator B which is the completion of $\mathcal{D} := \cap_{k=0}^{\infty}\mathcal{D}(B^k)$ with respect to the inner product $\langle u, v\rangle_r := \langle B^r u, B^r v\rangle$, $r \in \mathbf{R}$, $u, v \in \mathcal{D}$. We may observe that, for $x \in \mathcal{D}$, the associated Hilbert space norm is given by

$$\|x\|_r = \|B^r x\|, \qquad r \in \mathbf{R}.$$

According to [10], we call a function $\phi : \mathbf{R}^+ \to \mathbf{R}^+$ an *index function* if it is continuous and strictly increasing with $\phi(0+) = 0$.

Assumption A1. *For some $p > 0$ and $E < \infty$, the solution $x^\dagger$ of problem (1.1) is an element of the set*

$$M_{p,E} = \{x \in X \mid \|x\|_p \leqslant E\}. \tag{2.1}$$

Assumption A2. *There exists some index function ϕ with properties*

(i) there exists a constant $m > 0$ with

$$m\|[\phi(B^{-2})]^{1/2}x\| \leqslant \|Ax\| \qquad \textit{for all} \quad x \in X, \tag{2.2}$$

(ii) the function $\psi_p : (0, c] \to \mathbf{R}^+$ defined by $\psi_p(\lambda) := \lambda\phi(\lambda^{1/p})$ with $c = \|B^{-2p}\|$ and p as in assumption A1 is convex.

Assumption A1 characterizes the smoothness of the unknown solution $x^\dagger$ in the scale $(X_r)_{r\in\mathbf{R}}$. The operator B which defines the Hilbert scale is generally some differential operator such that B^{-1} is finitely smoothing. Hence, by using assumption A1 we can study different smoothness situations for the unknown solution $x^\dagger$ where in practice the parameter p characterizing the smoothness of $x^\dagger$ is generally unknown. Assumption A2 characterizes the relation between the smoothing properties of the operators A and B^{-1} in a general way allowing the study of finitely and infinitely smoothing operators A. In addition, the setting of this paper also allows us to consider the case that both A and B^{-1} are infinitely smoothing operators.

Note that inequality (2.2) implies the range inclusion $\mathcal{R}(G) \subset \mathcal{R}(|A|)$ with $G = [\phi(B^{-2})]^{1/2}$ and $|A| = (A^*A)^{1/2}$. By using such range inclusions, convergence rate results for method (1.3) with $s = 0$ have been obtained in [5]. Conversely, a range inclusion $\mathcal{R}(G) \subset \mathcal{R}(|A|)$ implies inequality (2.2) for some constant $m > 0$. Details and consequences of this fact may be found in paper [2].

Now we discuss the concepts of optimality and order optimality of an approximation method for problems (1.1) and (1.2). Any operator $R : Y \to X$ can be considered as a special method for solving problems (1.1) and (1.2). Thus, corresponding to inexact data y^δ satisfying (1.2), Ry^δ can be considered as an approximate solution to (1.1). Given a method R and the error level $\delta > 0$, the quantity

$$\Delta(\delta, R) = \sup\{\|Ry^\delta - x^\dagger\| \mid x^\dagger \in M_{p,E}, y^\delta \in Y, \|y - y^\delta\| \leqslant \delta\}$$

is the *worst case error* for identifying the solution $x^\dagger$ of problem (1.1) from $y^\delta \in Y$ under the assumptions (1.2) and $x^\dagger \in M_{p,E}$. This worst case error characterizes the maximal error of the method R if the solution $x^\dagger$ of problem (1.1) varies in the set $M_{p,E}$. An optimal method R_{opt} is characterized by $\Delta(\delta, R_{\mathrm{opt}}) = \inf_R \Delta(\delta, R)$. It is well known (see, e.g, [23, p 8]) that $\inf_R \Delta(\delta, R) \geqslant \omega(\delta, M_{p,E})$ where the *modulus of continuity* $\omega(\delta, M_{p,E})$ of the inverse operator A^{-1} on the set $M_{p,E}$ is given by

$$\omega(\delta, M_{p,E}) = \sup\{\|x\| \mid x \in M_{p,E}, \|Ax\| \leqslant \delta\}. \tag{2.3}$$

Moreover, we have $\inf_R \Delta(\delta, R) \leqslant 2\omega(\delta, M_{p,E})$ as well (see, e.g., [11]). In view of these relations, one may look for a regularization method R_α^δ, or for a regularized approximation $x_\alpha^\delta = R_\alpha^\delta y^\delta$ together with a parameter choice strategy $\alpha := \alpha(\delta, y^\delta)$ such that $\left\|x_\alpha^\delta - x^\dagger\right\| \leqslant \omega(\delta, M_{p,E})$, or at least

$$\left\|x_\alpha^\delta - x^\dagger\right\| \leqslant c\omega(\delta, M_{p,E}) \tag{2.4}$$

for some positive constant $c \geqslant 1$. For the case of $s = p$, order optimal estimates of the form (2.4) are available in the literature—see [9] and [13] for the Tikhonov regularization with an *a priori* and *a posteriori* choice of α, respectively, and [14] for a general regularization method. However, it is always desirable to obtain some order optimal estimate for $\left\|x_\alpha^\delta - x^\dagger\right\|$ in terms of δ so that the decay of the error can be inferred from the nature of the dependence of the estimate on δ. Thus, one would like to have some sharp estimate for the quantity $\omega(\delta, M_{p,E})$. Our next job is to do that under the additional assumption A2.

In our first proposition we provide some bound for elements x satisfying $\|x\|_p \leqslant 1$ and assumption A2. The proof of this proposition is along the line of Mair's paper [9]. For the sake completeness of exposition we include the proof.

Proposition 2.1. *Let $\|x\|_p \leqslant 1$ and let assumption A2 be satisfied. Then*

$$\|x\| \leqslant \sqrt{\psi_p^{-1}\left(\frac{\|Ax\|^2}{m^2}\right)}. \tag{2.5}$$

Proof. Let $\|x\|_p \leqslant 1$ and let E_λ be the spectral family of B^{2p}. Since ψ_p is convex, we may employ Jensen's inequality and obtain due to assumption A2 the estimate

$$\psi_p\left(\frac{\|x\|^2}{\|x\|_p^2}\right) \leqslant \frac{\int \psi_p\left(\frac{1}{\lambda}\right)\lambda \, \mathrm{d}\|E_\lambda x\|^2}{\|x\|_p^2} = \frac{\|[\phi(B^{-2})]^{1/2}x\|^2}{\|x\|_p^2} \leqslant \frac{\|Ax\|^2}{m^2\|x\|_p^2}.$$

From $\|x\|_p \leqslant 1$ we have $\|x\| \leqslant \|x\|/\|x\|_p$. Consequently, since ϕ is monotone,

$$\phi(\|x\|^{2/p}) \leqslant \phi\left(\frac{\|x\|^{2/p}}{\|x\|_p^{2/p}}\right) = \frac{\|x\|_p^2}{\|x\|^2} \cdot \psi_p\left(\frac{\|x\|^2}{\|x\|_p^2}\right) \leqslant \frac{\|Ax\|^2}{m^2\|x\|^2}.$$

We multiply this inequality by $\|x\|^2$ and obtain $\psi_p(\|x\|^2) \leqslant \|Ax\|^2/m^2$. From this estimate we obtain (2.5). □

For estimating the modulus of continuity $\omega(\delta, M_{p,E})$ of the inverse operator A^{-1} on the set $M_{p,E}$, we make use of assumption A2 and proposition 2.1.

Theorem 2.2. *Let $M_{p,E}$ be given by (2.1) and let assumption A2 be satisfied. Then,*

$$\omega(\delta, M_{p,E}) \leqslant E\sqrt{\psi_p^{-1}\left(\frac{\delta^2}{m^2E^2}\right)}. \tag{2.6}$$

If there holds equality in (2.2) and if $\delta_0 := \delta/(mE)$ is an element of the spectrum $\sigma(G)$ of the operator $G := [\phi(B^{-2})]^{1/2}B^{-p}$, then there holds equality in (2.6).

Proof. The estimate (2.6) follows from proposition 2.1 by taking x/E in place of x. Let us prove the second part of the theorem. If there holds equality in (2.2), then (2.3) attains the form

$$\begin{aligned}\omega(\delta, M_{p,E}) &= \sup\{\|x\| \mid \|[\phi(B^{-2})]^{1/2}x\| \leqslant \delta/m \wedge \|B^p x\| \leqslant E\}\\ &= \sup\{\|B^{-p}v\| \mid \|Gv\| \leqslant \delta/m \wedge \|v\| \leqslant E\}.\end{aligned} \tag{2.7}$$

Assume that $\delta_0 := \delta/(mE)$ is an eigenvalue of the operator $G = [\phi(B^{-2})]^{1/2}B^{-p}$ and v_0 the corresponding eigenelement with $\|v_0\| = E$; then we have

$$Gv_0 = \delta_0 v_0. \tag{2.8}$$

Consequently, $\|Gv_0\| = \delta/m$ and $\|v_0\| = E$. Hence, in view of (2.7) we conclude that

$$\omega(\delta, M_{p,E}) \geqslant \|B^{-p}v_0\|. \tag{2.9}$$

Exploiting the definition of ψ_p, from (2.8) we have $\psi_p(B^{-2p})v_0 = \delta_0^2 v_0$, or equivalently, $B^{-p}v_0 = \sqrt{\psi_p^{-1}(\delta_0^2)}\,v_0$. Hence, (2.9) provides $\omega(\delta, M_{p,E}) \geqslant E\sqrt{\psi_p^{-1}(\delta_0^2)}$, and due to (2.6) we have $\omega(\delta, M_{p,E}) = E\sqrt{\psi_p^{-1}(\delta_0^2)}$. If $\delta_0 \in \sigma(G)$ is not an eigenvalue, then δ_0 belongs to the continuous spectrum of G and the proof of the equality in (2.6) follows with small modifications. □

Let us discuss two special cases. For simplicity, in these two special cases inequality (2.2) of assumption A2 is satisfied as equality. Some more general examples in which the the smoothing properties of the operator A relative to the operator B^{-1} are characterized not by equality, but more generally, by some inequalities (2.2) and (4.1), will be discussed in section 6.

Example 2.3 (finitely smoothing case)**.** Let us assume that the operators A^*A and B are related by

$$A^*A = B^{-2a} \tag{2.10}$$

where a is some positive constant. Such situations occur, e.g., for numerical differentiation problems of certain order. In this special case, assumption A2 (i) holds true as equality with $m = 1$ and $\phi(\lambda) = \lambda^a$. We easily see that the function ϕ is an index function and that ψ_p attains the form $\psi_p(\lambda) = \lambda^{(a+p)/p}$ and satisfies A2 (ii). Computing the right-hand side of (2.6) we find that for example (2.10) the modulus of continuity $\omega(\delta, M_{p,E})$ of the inverse operator A^{-1} on the set $M_{p,E}$ is given by

$$\omega(\delta, M_{p,E}) = O\big(\delta^{\frac{p}{p+a}}\big).$$

Example 2.4 (infinitely smoothing case)**.** Let us assume that the operators A^*A and B are related by

$$A^*A = \mathrm{e}^{-B^a} \tag{2.11}$$

where a is some positive constant. Such situations occur, e.g., in inverse heat conduction problems. Here, assumption A2 (i) holds true with $m = 1, \phi(\lambda) = \mathrm{e}^{-\lambda^{-a/2}}$ and equality in (2.2). The function ψ_p takes the form $\psi_p(\lambda) = \lambda\,\mathrm{e}^{-\lambda^{-a/(2p)}}$, which is convex on the interval $(0, \|B^{-2p}\|]$ provided $\|B\| \geqslant 1$. Hence, if $\|B\| \geqslant 1$, then

$$\omega(\delta, M_{p,E}) = O([-\ln\delta]^{-p/a}).$$

Theorem 2.2 motivates the following definition (see, e.g., [23, p 8]).

Definition 2.5. *Let assumptions A1, A2 be satisfied. Then, a method R^δ, or the corresponding approximate solution $x^\delta = R^\delta y^\delta$, is called*

(i) optimal on the set $M_{p,E}$ if $\|x^\delta - x^\dagger\| \leqslant E\sqrt{\psi_p^{-1}\left(\frac{\delta^2}{m^2E^2}\right)}$,
(ii) order optimal on the set $M_{p,E}$ if $\|x^\delta - x^\dagger\| \leqslant cE\sqrt{\psi_p^{-1}\left(\frac{\delta^2}{m^2E^2}\right)}$ with $c \geqslant 1$.

3. Regularization

3.1. A general regularization scheme

Let us consider a general regularization scheme in Hilbert scales in which the regularized approximations with exact and noisy data y and y^δ, respectively, are defined by

$$x_\alpha = B^{-s} g_\alpha(T^*T)T^*y, \qquad x_\alpha^\delta = B^{-s} g_\alpha(T^*T)T^*y^\delta \qquad \text{with} \quad T = AB^{-s}. \tag{3.1}$$

Here, $s \geqslant 0$ is some nonnegative number that controls the smoothness to be introduced into the regularization process and $g_\alpha : (0, \|T\|^2] \to \mathbf{R}$ is a piecewise continuous function with the property that $\lim_{\alpha\to 0+} g_\alpha(\lambda) = 1/\lambda$. Different regularization methods are characterized by different functions g_α in (3.1). For the study of the general regularization method (3.1), besides assumptions A1 and A2 of section 2, the following additional assumption is required which is analogous to a corresponding assumption in [23].

Assumption A3. *There exist positive constants γ_1 and β_1 such that*

(i) $\sup_{\lambda>0} \lambda^{1/2}|g_\alpha(\lambda)| \leqslant \gamma_1/\sqrt{\alpha}$, $\quad \sup_{\lambda>0} \lambda|g_\alpha(\lambda)| \leqslant 1$,
(ii) $\sup_{\lambda>0} \lambda^{1/2}|1 - \lambda g_\alpha(\lambda)| \leqslant \beta_1\sqrt{\alpha}$, $\quad \sup_{\lambda>0} |1 - \lambda g_\alpha(\lambda)| \leqslant 1$.

Let us discuss some special regularization methods that fit into the framework of the general regularization scheme (3.1) and which satisfy assumption A3.

Example 3.1 (ordinary Tikhonov regularization in Hilbert scales)**.** This method is characterized by (3.1) with $g_\alpha(\lambda) = 1/(\lambda + \alpha)$. The regularized approximation x_α^δ can be obtained by solving the minimization problem

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

or as the solution of the associated operator equation $(A^*A + \alpha B^{2s})x_\alpha^\delta = A^*y^\delta$. In this example, assumption A3 is satisfied with constants $\gamma_1 = 1/2$ and $\beta_1 = 1/2$.

Example 3.2 (Tikhonov regularization of order m in Hilbert scales)**.** These methods are characterized by (3.1) with $g_\alpha(\lambda) = \left(1 - \left(\frac{\alpha}{\lambda+\alpha}\right)^m\right)/\lambda$. The regularized approximations $x_\alpha^\delta := x_{\alpha,m}^\delta$ can be obtained by solving the m linear operator equations

$$(A^*A + \alpha B^{2s})x_{\alpha,k}^\delta = A^*y^\delta + \alpha B^{2s}x_{\alpha,k-1}^\delta, \qquad k = 1, \ldots, m, \quad x_0^\delta = 0.$$

For $m = 1$, this method coincides with the method of example 3.1. For $m \geqslant 2$, assumption A3 is satisfied with constants $\gamma_1 = \sqrt{m}$ and $\beta_1 = 1$ (see [23]).

Example 3.3 (spectral method in Hilbert scales)**.** Consider method (3.1) with

$$g_\alpha(\lambda) = \begin{cases} 1/\lambda & \text{for} \quad \lambda \geqslant \alpha \\ 1/\alpha & \text{for} \quad \lambda \leqslant \alpha. \end{cases}$$

For problems with the compact operators A and B^{-1}, the numerical computation of x_α^δ can effectively be done by

$$x_\alpha^\delta = \sum_{s_i \geqslant \sqrt{\alpha}} \frac{\langle y^\delta, v_i \rangle}{s_i} u_i + \frac{1}{\alpha} \left(B^{-2s} A^* y^\delta - \sum_{s_i \geqslant \sqrt{\alpha}} s_i \langle y^\delta, v_i \rangle u_i \right),$$

which requires the computation of only *finite* sums. Here $\{s_i, u_i, v_i\}_{i\in\mathbf{N}}$ denotes the generalized singular system of A satisfying $A^*Au_i = \lambda_i B^{2s} u_i$, $s_i = \sqrt{\lambda_i}$ and $v_i = \frac{1}{s_i} A u_i$. In fact, $\{s_i, u_i, v_i\}_{i\in\mathbf{N}}$ is a singular system for the compact operator $T = AB^{-s}$. For this method, assumption A3 holds true with $\gamma_1 = 1$ and $\beta_1 = 2/\sqrt{27}$ (see [23]).

Example 3.4 (asymptotical regularization in Hilbert scales). This method is characterized by (3.1) with $g_\alpha(\lambda) = (1 - \mathrm{e}^{-\lambda/\alpha})/\lambda$. In this method one solves the Cauchy problem

$$B^{2s}\dot{u}(t) + A^*Au(t) = A^*y^\delta, \qquad 0 < t \leqslant \tau, \quad u(0) = 0$$

and the regularized approximation is defined by $x_\alpha^\delta = u(\tau)$. Here τ and α are related by $\tau = 1/\alpha$. For this regularization method assumption A3 is satisfied with constants $\gamma_1 = 1$ and $\beta_1 = 1/\sqrt{2\mathrm{e}}$ (see [23]).

Example 3.5 (iterative regularization in Hilbert scales). As a special case of more general iterative regularization methods, let us consider the *Landweber iteration*. This method is characterized by (3.1) with $g_\alpha(\lambda) = (1 - (1-\lambda)^{1/\alpha})/\lambda$. The regularized approximation $x_\alpha^\delta := u_n^\delta$ can be obtained by performing n iterations according to

$$u_k^\delta = u_{k-1}^\delta - B^{-2s}A^*\big(Au_{k-1}^\delta - y^\delta\big), \qquad k = 1, \ldots, n,$$

with $u_0 = 0$. Here, n and α are related by $\alpha = 1/n$. For this regularization method assumption A3 is satisfied with constants $\gamma_1 = 1$ and $\beta_1 = 1/\sqrt{2\mathrm{e}}$ (see [23]).

3.2. A priori parameter choice

In this subsection we will prove that under assumption A2 the regularized approximation x_α^δ from (3.1) with $s = p$ is order optimal on the set $M_{p,E}$ provided α is chosen *a priori*. From this result we deduce, as a special case, Mair's convergence rate result for the method of Tikhonov regularization (see [9]).

Theorem 3.6. *Let x_α^δ be the regularized approximation (3.1) with s chosen by $s = p$ and let assumptions A1 and A3 be satisfied. Then, for $\alpha = \delta^2/E^2$,*

$$\big\|x_\alpha^\delta - x^\dagger\big\| \leqslant (\gamma_1 + 1)\omega(c\delta, M_{p,E}) \qquad \text{with} \quad c = \frac{\beta_1 + 1}{\gamma_1 + 1}. \tag{3.2}$$

If, in addition, assumption A2 is satisfied, then

$$\big\|x_\alpha^\delta - x^\dagger\big\| \leqslant (\gamma_1 + 1)E\sqrt{\psi_p^{-1}\left(\frac{c^2\delta^2}{m^2E^2}\right)}. \tag{3.3}$$

Proof. Due to (3.1) there hold the representations

$$\begin{aligned} A\big(x_\alpha^\delta - x_\alpha\big) &= Tg_\alpha(T^*T)T^*(y^\delta - y) \\ A(x^\dagger - x_\alpha) &= T[I - g_\alpha(T^*T)T^*T]B^s x^\dagger \\ B^s\big(x_\alpha^\delta - x_\alpha\big) &= g_\alpha(T^*T)T^*(y^\delta - y) \\ B^s(x^\dagger - x_\alpha) &= [I - g_\alpha(T^*T)T^*T]B^s x^\dagger \end{aligned} \tag{3.4}$$

with $T = AB^{-s}$. From (3.4) with $s = p$, assumptions A1 and A3, the triangle inequality and the parameter choice $\alpha = \delta^2/E^2$, we obtain the two estimates

$$\|Ax_\alpha^\delta - Ax^\dagger\| \leqslant \delta + \beta_1\sqrt{\alpha}E = (\beta_1 + 1)\delta, \qquad \|x_\alpha^\delta - x^\dagger\|_p \leqslant \gamma_1\delta/\sqrt{\alpha} + E = (\gamma_1 + 1)E.$$

Hence, (3.2) follows. Now (3.3) is a consequence of theorem 2.2. □

Remark 3.7. Note that by taking $k = \max\{\beta_1, \gamma_1\}$, the last two estimates in the proof of the above theorem imply $\|Ax_\alpha^\delta - Ax^\dagger\| \leqslant (k+1)\delta$ and $\|x_\alpha^\delta - x^\dagger\|_p \leqslant (k+1)E$. Hence, instead of (3.2) and (3.3) we have the two estimates

$$\|x_\alpha^\delta - x^\dagger\| \leqslant (k+1)\omega(\delta, M_{p,E}), \qquad \|x_\alpha^\delta - x^\dagger\| \leqslant (k+1)E\sqrt{\psi_p^{-1}\left(\frac{\delta^2}{m^2E^2}\right)}, \tag{3.5}$$

respectively. In fact, the second error bound of (3.5) shows the order optimality of the regularized approximation x_α^δ in the sense of definition 2.5. The second error bound of (3.5) can also be derived from (3.3) by making use of the relation

$$\psi_p^{-1}(c\lambda) \leqslant c_1\psi_p^{-1}(\lambda) \qquad \text{with} \quad c_1 = \max\{1, c\} \tag{3.6}$$

that can be realized as follows. Due to the monotonicity of ϕ we have for arbitrary $p > 0$ the estimate $ct\phi(t^{1/p}) \leqslant c_1t\phi((c_1t)^{1/p})$, or equivalently, $c\psi_p(t) \leqslant \psi_p(c_1t)$, since $\psi_p(\lambda) = \lambda\phi(\lambda^{1/p})$. Since ϕ is an index function it follows that ψ_p and hence ψ_p^{-1} is strictly monotonically increasing. Consequently, $\psi_p^{-1}(c\psi_p(t)) \leqslant c_1t$. Now, choosing $t = \psi_p^{-1}(\lambda)$ we obtain (3.6).

3.3. Discrepancy principle

In this subsection we study the case of choosing the regularization parameter α *a posteriori* by Morozov's discrepancy principle.

Morozov's discrepancy principle. For a given constant $C \geqslant 1$, choose α as the solution of the nonlinear scalar equation

$$d(\alpha) := \|Ax_\alpha^\delta - y^\delta\| = C\delta. \tag{3.7}$$

For guaranteeing that equation (3.7) possesses a unique solution one has to assume that the function $g_\alpha : (0, a] \to (0, \infty)$ with $a \geqslant \|T\|^2$ and $\alpha > 0$ satisfies the following:

(i) $\sup_{0\leqslant\lambda\leqslant a}|1 - \lambda g_\alpha(\lambda)| \leqslant 1$ and $\sup_{0\leqslant\lambda\leqslant a}|g_\alpha(\lambda)| \leqslant \gamma/\alpha$ for a constant $\gamma > 0$,
(ii) $1 - \lambda g_\alpha(\lambda) \to 0$ for $\alpha \to 0$ and all $\lambda \in [0, a]$,
(iii) $|1 - \lambda g_{\alpha_1}(\lambda)| \leqslant |1 - \lambda g_{\alpha_2}(\lambda)|$ for $\alpha_1 \leqslant \alpha_2$,
(iv) $g_{\alpha_n}(\lambda) \to g_\alpha(\lambda)$ for $\alpha_n \to \alpha > 0$ and all $\lambda \in [0, a]$.

Since A is injective, from [23, p 64, lemma 3.1] we have that under the above conditions (i) and (ii) there hold the limit relations

$$\lim_{\alpha\to 0} d(\alpha) = 0 \quad \text{and} \quad \lim_{\alpha\to\infty} d(\alpha) = \|y^\delta\|.$$

In addition, we have from [23, p 64, lemma 3.1] that under the conditions (iii) and (iv) the function d is monotonically increasing and continuous. Hence, under the above conditions (i)–(iv) the nonlinear scalar equation (3.7) possesses a unique solution provided $C\delta < \|y^\delta\|$. We note that the above conditions (i)–(iv) are satisfied for the regularization methods discussed in examples 3.1–3.5. Hence, in working with the discrepancy principle (3.7) we will always assume that the above conditions (i)–(iv) are satisfied without mentioning it explicitly.

For deriving order optimal error bounds for $\|x_\alpha^\delta - x^\dagger\|$ with α chosen according to the discrepancy principle (3.7), we exploit the following assumption from [23, p 75]:

Assumption A4. *The function* $g_\alpha : (0, a] \to (0, \infty)$ *with* $a = \|T\|^2$ *and* $\alpha > 0$ *satisfies*

$$\alpha g_\alpha(\lambda) \leqslant 1 \qquad \text{and} \qquad 0 \leqslant 1 - \lambda g_\alpha(\lambda) \leqslant \alpha g_\alpha(\lambda).$$

Note that for the methods discussed in examples 3.1–3.5 assumption A4 is satisfied. For providing order optimal error bounds for $\|x_\alpha^\delta - x^\dagger\|$ the following proposition is useful which holds true under assumption A4 and can be proved along the line of the proof of [23, p 77, lemma 4.1], where the special case $s = 0$ has been treated.

Proposition 3.8. *Let* x_α^δ *be the regularized approximation (3.1) with* $s \geqslant 0$ *and let assumption A4 be satisfied. Then,*

$$\|Ax_\alpha^\delta - y^\delta\|^2 + \alpha\|x_\alpha^\delta - x^\dagger\|_s^2 \leqslant \|y - y^\delta\|^2 + \alpha\left\|[I - T^*Tg_\alpha(T^*T)]^{\frac{1}{2}}B^s x^\dagger\right\|^2. \tag{3.8}$$

Proposition 3.8 is the basic ingredient for deriving order optimal error bounds for $\|x_\alpha^\delta - x^\dagger\|$ in case α is chosen by the discrepancy principle, not only in the special case $s = p$, but also for the practically more important low-order case $s < p$ in section 5.

Theorem 3.9. *Let* x_α^δ *be the regularized approximation from (3.1) with s chosen by* $s = p$, *let assumptions A1 and A4 be satisfied and let* α *be chosen according to the discrepancy principle (3.7). Then*

$$\|x_\alpha^\delta - x^\dagger\| \leqslant \omega(c\delta, M_{p,E}) \qquad \text{with} \quad c = C + 1. \tag{3.9}$$

If in addition assumption A2 is satisfied, then

$$\|x_\alpha^\delta - x^\dagger\| \leqslant E\sqrt{\psi_p^{-1}\left(\frac{c^2\delta^2}{m^2E^2}\right)}. \tag{3.10}$$

Proof. For α chosen by the discrepancy principle (3.7) the estimate (3.8) with $s = p$ attains the form $C^2\delta^2 + \alpha\|x_\alpha^\delta - x^\dagger\|_p^2 \leqslant \delta^2 + \alpha\|R_\alpha^{1/2}B^p x^\dagger\|^2$ with $R_\alpha = I - T^*Tg_\alpha(T^*T)$. Since $C \geqslant 1$ we have $\|x_\alpha^\delta - x^\dagger\|_p \leqslant \|R_\alpha^{1/2}B^p x^\dagger\| \leqslant \|R_\alpha^{1/2}\| \cdot \|B^p x^\dagger\|$. Due to assumption A4 we have $\|R_\alpha^{1/2}\| \leqslant 1$. We exploit in addition assumption A1 and obtain $\|x_\alpha^\delta - x^\dagger\|_p \leqslant E$. In addition, the discrepancy principle (3.7) and the triangle inequality provide

$$\|Ax_\alpha^\delta - Ax^\dagger\| \leqslant \|Ax_\alpha^\delta - y^\delta\| + \|y - y^\delta\| \leqslant (C+1)\delta.$$

From this estimate and $\|x_\alpha^\delta - x^\dagger\|_p \leqslant E$ we obtain (3.9). Now (3.10) is a consequence of theorem 2.2. □

4. High-order regularization

In this section, we study the case of high-order regularization in which the parameter s in methods (3.1) is larger than the number p in assumption A1 characterizing the smoothness of the unknown solution $x^\dagger$. We will prove that in the case $s \geqslant p$ the same order optimal error bounds $\|x_\alpha^\delta - x^\dagger\| = O\left(\sqrt{\psi_p^{-1}(\delta^2)}\right)$ can be guaranteed as in the case $s = p$. This means, in particular, that there is no loss of accuracy if s is chosen larger than p.

4.1. A priori parameter choice

In our first subsection we study the case of *a priori* parameter choice. As in the foregoing section 3 we exploit assumption A1 and replace assumption A2 by a stronger assumption in which assumption A2 (i) is replaced by a two-sided estimate and assumption A2 (ii) holds true with p replaced by s.

Assumption A5. *There exists some index function ϕ with properties*

(i) there exist positive constants m and M with

$$m\left\|[\phi(B^{-2})]^{1/2}x\right\| \leqslant \|Ax\| \leqslant M\left\|[\phi(B^{-2})]^{1/2}x\right\| \qquad \textit{for all} \quad x \in X, \tag{4.1}$$

(ii) for $s > 0$, the function $\psi_s : (0, c] \to \mathbf{R}^+$ defined by $\psi_s(\lambda) := \lambda\phi(\lambda^{1/s})$ with $c = \|B^{-2s}\|$ and s as in (3.1) is convex.

The way of deriving order optimal error bounds is borrowed from [9] where the special case of example 3.1 has been treated and consists in constructing some sufficiently smooth approximation $x_0 \in X_s$ for $x^\dagger \in X_p$ such that both the error parts $\left\|x_\alpha^\delta - x_0\right\|$ and $\|x_0 - x^\dagger\|$ are of the same order $O\big(\sqrt{\psi_p^{-1}(\delta^2)}\big)$. We construct x_0 according to

$$x_0 = \int_b^\tau \mathrm{d}E_\lambda x^\dagger \tag{4.2}$$

where E_λ is the spectral measure of B, $b = 1/\|B^{-1}\|$ and $\tau < \infty$ has to be chosen properly. From [9] we have the following.

Proposition 4.1. *Let $\|x^\dagger\|_p \leqslant E$ and let $x_0 \in X_s$ be given by (4.2). Then, for $s \geqslant p$,*

$$\|x^\dagger - x_0\| \leqslant \tau^{-p}E \qquad \textit{and} \qquad \|x_0\|_s \leqslant \tau^{s-p}E. \tag{4.3}$$

Let, in addition, assumption A5 be satisfied. Then

$$\|Ax^\dagger - Ax_0\| \leqslant ME\sqrt{\psi_p(\tau^{-2p})}. \tag{4.4}$$

In our further studies, we shall make use of the the following result.

Proposition 4.2. *Let ψ_s be defined as in assumption A5. Then, for arbitrary positive constants s, p, c,*

$$\psi_s^{-1}\big(c\delta^2\big[\psi_p^{-1}(\delta^2)\big]^{(s-p)/p}\big) \leqslant c_1\big[\psi_p^{-1}(\delta^2)\big]^{s/p} \tag{4.5}$$

with $c_1 = \max\{1, c\}$. If $c = 1$, then there holds equality in (4.5).

From $\psi_p(\lambda) = \lambda\phi(\lambda^{1/p})$ and $\psi_s(\lambda^{s/p}) = \lambda^{s/p}\phi(\lambda^{1/p})$ we obtain the identity $\psi_p(\lambda)\cdot\lambda^{(s-p)/p} = \psi_s(\lambda^{s/p})$. We substitute $\lambda = \psi_p^{-1}(\delta^2)$, multiply by c, apply on both sides ψ_s^{-1} and obtain

$$\psi_s^{-1}\big(c\delta^2\big[\psi_p^{-1}(\delta^2)\big]^{(s-p)/p}\big) = \psi_s^{-1}\big(c\psi_s\big(\big[\psi_p^{-1}(\delta^2)\big]^{s/p}\big)\big).$$

Hence, for $c = 1$ we have equality in (4.5). Now, in view of the relation (3.6), we have $\psi_s^{-1}(c\lambda) \leqslant c_1\psi_s^{-1}(\lambda)$ with $\lambda = \psi_s\big(\big[\psi_p^{-1}(\delta^2)\big]^{s/p}\big)$, so that we obtain (4.5).

Now we are ready to provide order optimal error bounds for regularized approximations (3.1) under certain *a priori* choice of the regularization parameter.

Theorem 4.3. *Let x_α^δ be the regularized approximation as in (3.1) and let assumptions A1, A3 and A5 be satisfied. If α is chosen a priori by*

$$\alpha = \delta^2\big[\psi_p^{-1}(\delta^2)\big]^{(s-p)/p}, \tag{4.6}$$

then, for $s \geqslant p$,

$$\left\|x_\alpha^\delta - x^\dagger\right\| = O\Big(\sqrt{\psi_p^{-1}(\delta^2)}\Big). \tag{4.7}$$

Proof. Let x_0 be defined by (4.2) and define $x_{\alpha,0}$ by $x_{\alpha,0} = B^{-s} g_\alpha(T^*T)T^*Ax_0$ with $T = AB^{-s}$. Then, in analogy to (3.4), we have

$$\begin{aligned}
&A\big(x_\alpha^\delta - x_{\alpha,0}\big) = Tg_\alpha(T^*T)T^*(y^\delta - Ax_0),\\
&A(x_0 - x_{\alpha,0}) = T[I - g_\alpha(T^*T)T^*T]B^s x_0,\\
&B^s\big(x_\alpha^\delta - x_{\alpha,0}\big) = g_\alpha(T^*T)T^*(y^\delta - Ax_0),\\
&B^s(x_0 - x_{\alpha,0}) = [I - g_\alpha(T^*T)T^*T]B^s x_0.
\end{aligned} \tag{4.8}$$

From (4.8), assumption A3, proposition 4.1 and the triangle inequality, we obtain

$$\begin{aligned}
\left\|Ax_\alpha^\delta - Ax_0\right\| &\leqslant \left\|A\big(x_\alpha^\delta - x_{\alpha,0}\big)\right\| + \|A(x_0 - x_{\alpha,0})\|\\
&\leqslant \delta + ME\sqrt{\psi_p(\tau^{-2p})} + \beta_1\sqrt{\alpha}\,\tau^{s-p}E
\end{aligned} \tag{4.9}$$

and

$$\left\|x_\alpha^\delta - x_0\right\|_s \leqslant \frac{\gamma_1}{\sqrt{\alpha}}\Big(\delta + ME\sqrt{\psi_p(\tau^{-2p})}\Big) + \tau^{s-p}E. \tag{4.10}$$

Now we choose τ such that $\delta = \sqrt{\psi_p(\tau^{-2p})}$ holds, which gives

$$\tau = \big[\psi_p^{-1}(\delta^2)\big]^{-1/2p}. \tag{4.11}$$

The parameter choice (4.6) is equivalent to $\tau^{s-p} = \delta/\sqrt{\alpha}$. Hence, from (4.9) and (4.10) we have

$$\left\|Ax_\alpha^\delta - Ax_0\right\| \leqslant c_1\delta \qquad \text{and} \qquad \left\|x_\alpha^\delta - x_0\right\|_s \leqslant c_2\big[\psi_p^{-1}(\delta^2)\big]^{(p-s)/(2p)} \tag{4.12}$$

with the two constants $c_1 := 1 + ME + \beta_1 E$ and $c_2 := \gamma_1(1 + ME) + E$. Thus, we have $\left\|x_\alpha^\delta - x_0\right\| \leqslant \omega(c_1\delta, M_{s,E_1})$ with $E_1 = c_2\big[\psi_p^{-1}(\delta^2)\big]^{(p-s)/(2p)}$, so that by theorem 2.2 and proposition 4.2 we obtain

$$\begin{aligned}
\left\|x_\alpha^\delta - x_0\right\| &\leqslant E_1\sqrt{\psi_s^{-1}\left(\frac{c_1^2\delta^2}{m^2E_1^2}\right)}\\
&= c_2\big[\psi_p^{-1}(\delta^2)\big]^{(p-s)/(2p)}\sqrt{\psi_s^{-1}\left(\frac{c_1^2\delta^2}{m^2c_2^2}\big[\psi_p^{-1}(\delta^2)\big]^{(s-p)/p}\right)}\\
&\leqslant c_3\sqrt{\psi_p^{-1}(\delta^2)}
\end{aligned} \tag{4.13}$$

with $c_3 = \max\{c_2, c_1/m\}$. Now (4.7) follows from (4.13), the first estimate of (4.3) with τ chosen by (4.11) and the triangle inequality. □

4.2. Discrepancy principle

In this subsection we provide order optimal error bounds for methods (3.1) in case the regularization parameter α is chosen *a posteriori* by Morozov's discrepancy principle (3.7). We start by providing a lower bound for the regularization parameter α obtained by (3.7) in terms of the data error δ.

Proposition 4.4. *Let assumptions A1, A3 (i) and A5 be satisfied and let x_α^δ be defined by (3.1) with $s \geqslant p$. If α is chosen by (3.7) with $C > 1$, then*

$$\alpha \geqslant \delta_1^2\big[\psi_p^{-1}\big(\delta_1^2\big)\big]^{(s-p)/p} \qquad \textit{with} \quad \delta_1 = \frac{(C-1)\delta}{2E\max\{M,\beta_1\}}. \tag{4.14}$$

Proof. Let $R_\alpha = I - g_\alpha(TT^*)TT^*$, let x_0 be given by (4.2) and let α be chosen by (3.7) with $C > 1$. We use the representation $y^\delta - Ax_\alpha^\delta = R_\alpha y^\delta$ and obtain from assumptions A3 (i), A5 and proposition 4.1

$$\begin{aligned} C\delta &= \big\|Ax_\alpha^\delta - y^\delta\big\| \\ &\leqslant \|R_\alpha(y^\delta - y)\| + \|R_\alpha(y - Ax_0)\| + \|R_\alpha Ax_0\| \\ &\leqslant \delta + \|y - Ax_0\| + \|R_\alpha T\| \cdot \|B^s x_0\| \\ &\leqslant \delta + ME\sqrt{\psi_p(\tau^{-2p})} + \beta_1 E\sqrt{\alpha}\tau^{s-p} \\ &\leqslant \delta + E\max\{M,\beta_1\}\Big[\sqrt{\psi_p(\tau^{-2p})} + \sqrt{\alpha}\tau^{s-p}\Big]. \end{aligned} \tag{4.15}$$

We choose τ as the solution of the equation $\sqrt{\psi_p(\tau^{-2p})} = \sqrt{\alpha}\tau^{s-p}$, that is, as the solution of $\tau^{2p-2s}\psi_p(\tau^{-2p}) = \alpha$, or equivalently, $\psi_s(\tau^{-2s}) = \alpha$, which gives $\tau = \big[\psi_s^{-1}(\alpha)\big]^{-1/(2s)}$. For this choice of τ the estimate (4.15) may be rewritten as $\delta_1 \leqslant \sqrt{\psi_p(\tau^{-2p})}$, or equivalently, as $\psi_p\big(\delta_1^2\big) \leqslant \tau^{-2p}$, which gives $\psi_p\big(\delta_1^2\big) \leqslant \big[\psi_s^{-1}(\alpha)\big]^{p/s}$, or equivalently,

$$\alpha \geqslant \psi_s\big(\big[\psi_p^{-1}\big(\delta_1^2\big)\big]^{s/p}\big). \tag{4.16}$$

From proposition 4.2 we know that $\psi_s\big(\big[\psi_p^{-1}\big(\delta_1^2\big)\big]^{s/p}\big) = \delta_1^2\big[\psi_p^{-1}\big(\delta_1^2\big)\big]^{(s-p)/p}$. Hence, (4.14) follows from (4.16). □

Theorem 4.5. *Let assumptions A1, A3 and A5 be satisfied and let x_α^δ be defined by (3.1) with $s \geqslant p$. If α is chosen by the discrepancy principle (3.7) with $C > 1$, then*

$$\big\|x_\alpha^\delta - x^\dagger\big\| = O\Big(\sqrt{\psi_p^{-1}(\delta^2)}\Big). \tag{4.17}$$

Proof. Let us define x_0 according to (4.2) with τ chosen by (4.11). By using (3.7), (4.4) and the fact that $\psi_p(\tau^{-2p}) = \psi_p\big(\psi_p^{-1}(\delta^2)\big) = \delta^2$, we obtain

$$\begin{aligned} \big\|Ax_\alpha^\delta - Ax_0\big\| &\leqslant \big\|Ax_\alpha^\delta - y^\delta\big\| + \|y - y^\delta\| + \|y - Ax_0\| \\ &\leqslant (C+1)\delta + ME\sqrt{\psi_p(\tau^{-2p})} \\ &\leqslant k_1\delta \end{aligned} \tag{4.18}$$

with $k_1 = C + 1 + ME$. From (4.10), the identity $\psi_p(\tau^{-2p}) = \delta^2$, (4.11), (4.14) and (3.6) we obtain for α chosen by (3.7)

$$\begin{aligned} \big\|x_\alpha^\delta - x_0\big\|_s &\leqslant \frac{\gamma_1(1+ME)\delta}{\sqrt{\alpha}} + E\big[\psi_p^{-1}(\delta^2)\big]^{(p-s)/(2p)} \\ &\leqslant \gamma_1(1+ME)\frac{\delta}{\delta_1}\big[\psi_p^{-1}\big(\delta_1^2\big)\big]^{(p-s)/(2p)} + E\big[\psi_p^{-1}(\delta^2)\big]^{(p-s)/(2p)} \\ &\leqslant k_2\big[\psi_p^{-1}(\delta^2)\big]^{(p-s)/(2p)} \end{aligned} \tag{4.19}$$

with some constant k_2. We proceed as in (4.13) and obtain from (4.18) and (4.19) that

$$\big\|x_\alpha^\delta - x_0\big\| \leqslant k_3\sqrt{\psi_p^{-1}(\delta^2)} \tag{4.20}$$

with $k_3 = \max\{k_2, k_1/m\}$. Finally, (4.17) follows from (4.20), the first estimate of (4.3) with τ chosen by (4.11) and the triangle inequality. □

5. Low-order regularization

In this section we study the case of low-order regularization in which the parameter s in methods (3.1) is smaller than the number p in assumption A1 characterizing the smoothness of the unknown solution $x^\dagger$. We will prove that in this case the same order optimal error bound $\left\|x_\alpha^\delta - x^\dagger\right\| = O\big(\sqrt{\psi_p^{-1}(\delta^2)}\big)$ can be guaranteed as in the case $s = p$. This means in particular that there is no loss of accuracy if the parameter s in method (1.3) is chosen smaller than p. To our knowledge, until now this result is only known for the finitely smoothing case discussed in section 1 as an example 2.3 (cf [20]). Our studies in case of low-order regularization do not require the two-sided estimate in assumption A5 (i). Instead we exploit the following additional assumption:

Assumption A6. *The function $f : (0, c] \to \mathbf{R}^+$ defined by*

$$f(\lambda) := \lambda^{s/(p-s)}\phi(\lambda^{1/(p-s)})$$

is convex, where $c := \|B^{-2p}\|$ and ϕ is the index function from assumption A2.

For example 2.3 discussed in section 1 the function f in assumption A6 attains the form $f(\lambda) = \lambda^{(a+s)/(p-s)}$; hence, assumption A6 holds true in the low-order case $s < p$ provided $s \geqslant (p-a)/2$. This is Natterer's side condition for proving (1.5).

5.1. A priori parameter choice

In our first proposition we estimate the regularization error in case of exact data.

Proposition 5.1. *Let x_α be the regularized approximation from (3.1) and let assumptions A1, A2, A5 (ii) and A6 hold. Assume further that*

$$\sup_{\lambda\geqslant 0}|1-\lambda g_\alpha(\lambda)| \leqslant 1 \qquad \textit{and} \qquad \sup_{\lambda\geqslant 0}\lambda|1-\lambda g_\alpha(\lambda)| \leqslant \beta_2\alpha \tag{5.1}$$

with some constant $\beta_2 > 0$. Then, for $0 \leqslant s < p$,

$$\|x_\alpha - x^\dagger\| \leqslant \begin{cases} E\left[\phi^{-1}\left(\dfrac{\beta_2\alpha}{m^2}\right)\right]^{p/2} & \textit{for} \quad s = 0 \\ E\left[\psi_s^{-1}\left(\dfrac{\beta_2\alpha}{m^2}\right)\right]^{p/(2s)} & \textit{for} \quad 0 < s < p. \end{cases} \tag{5.2}$$

Proof. Let us introduce the abbreviations $z_\alpha = x^\dagger - x_\alpha$ and $R_\alpha = I - g_\alpha(T^*T)T^*T$. From (3.4) we have the identity $B^s z_\alpha = R_\alpha B^s x^\dagger$, and due to the first inequality of (5.1) we have $\left\|R_\alpha^{1/2}\right\| \leqslant 1$. We use these properties and obtain due to the Cauchy Schwarz inequality and assumption A1 that

$$\|z_\alpha\|_s^2 = \|R_\alpha B^s x^\dagger\|^2 \leqslant \left\|R_\alpha^{1/2}B^s x^\dagger\right\|^2 = \langle B^{2s-p}z_\alpha, B^p x^\dagger\rangle \leqslant E\|z_\alpha\|_{2s-p}. \tag{5.3}$$

From (3.4) we have $Az_\alpha = TR_\alpha B^s x^\dagger$, and due to (5.1) we have $\left\|(T^*T)^{1/2}R_\alpha^{1/2}\right\|^2 \leqslant \beta_2\alpha$. We use these properties and obtain by using (5.3)

$$\|Az_\alpha\|^2 = \|(T^*T)^{1/2}R_\alpha B^s x^\dagger\|^2 \leqslant \beta_2\alpha\left\|R_\alpha^{1/2}B^s x^\dagger\right\|^2 \leqslant \beta_2\alpha E\|z_\alpha\|_{2s-p}. \tag{5.4}$$

Our next aim consists in deriving a third estimate that relates the three quantities $\|z_\alpha\|_s$, $\|Az_\alpha\|$ and $\|z_\alpha\|_{2s-p}$. Since f is convex, we may employ Jensen's inequality and have

$$f\left(\frac{\|z_\alpha\|_{2s-p}^2}{\|z_\alpha\|_s^2}\right) = f\left(\frac{\int\lambda^{s-p}\cdot\lambda^s\,\mathrm{d}\|E_\lambda z_\alpha\|^2}{\int\lambda^s\,\mathrm{d}\|E_\lambda z_\alpha\|^2}\right) \leqslant \frac{\int f(\lambda^{s-p})\cdot\lambda^s\,\mathrm{d}\|E_\lambda z_\alpha\|^2}{\int\lambda^s\,\mathrm{d}\|E_\lambda z_\alpha\|^2},$$

where E_λ is the spectral family of B^2. Since $f(\lambda^{s-p})\lambda^s = \phi(\lambda^{-1})$ we obtain by using assumption A2 that

$$f\left(\frac{\|z_\alpha\|_{2s-p}^2}{\|z_\alpha\|_s^2}\right) \leqslant \frac{\int \phi(\lambda^{-1})\,\mathrm{d}\|E_\lambda z_\alpha\|^2}{\|z_\alpha\|_s^2} = \frac{\|[\phi(B^{-2})]^{1/2} z_\alpha\|^2}{\|z_\alpha\|_s^2} \leqslant \frac{\|Az_\alpha\|^2}{m^2\|z_\alpha\|_s^2}. \tag{5.5}$$

For proper combining of the three estimates (5.3), (5.4) and (5.5) we introduce a new function g by $g(\lambda) := f(\lambda^2)/\lambda^2$. Since f is convex, we conclude that g is monotonically increasing. Hence, by (5.3), which may be rewritten as $\|z_\alpha\|_{2s-p}^{1/2}/E^{1/2} \leqslant \|z_\alpha\|_{2s-p}/\|z_\alpha\|_s$, the monotonicity of g and (5.5),

$$g\left(\frac{\|z_\alpha\|_{2s-p}^{1/2}}{E^{1/2}}\right) \leqslant g\left(\frac{\|z_\alpha\|_{2s-p}}{\|z_\alpha\|_s}\right) = \frac{\|z_\alpha\|_s^2}{\|z_\alpha\|_{2s-p}^2} f\left(\frac{\|z_\alpha\|_{2s-p}^2}{\|z_\alpha\|_s^2}\right) \leqslant \frac{\|Az_\alpha\|^2}{m^2\|z_\alpha\|_{2s-p}^2}.$$

Multiplying by $\|z_\alpha\|_{2s-p}/E$ and exploiting (5.4) yields

$$f\left(\frac{\|z_\alpha\|_{2s-p}}{E}\right) \leqslant \frac{\|Az_\alpha\|^2}{m^2 E\|z_\alpha\|_{2s-p}} \leqslant \frac{\beta_2\alpha}{m^2} \equiv \alpha_0. \tag{5.6}$$

Now we distinguish two cases $s = 0$ and $0 < s < p$. In the first case $s = 0$ the function f of assumption A6 attains the form $f(\lambda) = \phi(\lambda^{1/p})$. Consequently, $f^{-1}(\lambda) = [\phi^{-1}(\lambda)]^p$. By using this representation, the first estimate of (5.2) follows from (5.3) and (5.6). In the second case $0 < s < p$ we use the identity $f(\lambda^{(p-s)/s}) = \lambda\phi(\lambda^{1/s}) = \psi_s(\lambda)$ and obtain that the inverse f^{-1} possesses the representation $f^{-1}(\lambda) = \left[\psi_s^{-1}(\lambda)\right]^{(p-s)/s}$. We use this representation, combine (5.3), (5.6) and (5.4), (5.6), respectively, and obtain the two estimates

$$\|z_\alpha\|_s \leqslant E\left[\psi_s^{-1}(\alpha_0)\right]^{(p-s)/(2s)} \qquad \text{and} \qquad \|Az_\alpha\| \leqslant \sqrt{\beta_2\alpha}\,E\left[\psi_s^{-1}(\alpha_0)\right]^{(p-s)/(2s)}.$$

Now we proceed as in the proof of proposition 2.1 with $p := s$ and obtain

$$\begin{aligned}\|z_\alpha\| &\leqslant E\left[\psi_s^{-1}(\alpha_0)\right]^{(p-s)/(2s)} \sup\{\|x\| \mid \|Ax\| \leqslant \sqrt{\beta_2\alpha} \wedge \|x\|_s \leqslant 1\}\\ &\leqslant E\left[\psi_s^{-1}(\alpha_0)\right]^{(p-s)/(2s)}\sqrt{\psi_s^{-1}(\alpha_0)}\end{aligned}$$

which provides the second estimate of (5.2). □

Note that assumption (5.1) is satisfied for the special regularization methods discussed in examples 3.1–3.5. In fact, $\beta_2 = 1$ for example 3.1, $\beta_2 = (m-1)^{m-1}/m^m$ for example 3.2, $\beta_2 = 1/4$ for example 3.3, $\beta_2 = 1/\mathrm{e}$ for example 3.4 and $\beta_2 = 1/\mathrm{e}$ for example 3.5 (see [23]).

Now we are ready to provide order optimal error bounds for the total error $\left\|x_\alpha^\delta - x^\dagger\right\|$ in case of proper *a priori* parameter choice.

Theorem 5.2. *Let x_α^δ be the regularized approximation (3.1) and let the assumptions of proposition 5.1 and A3 (i) be satisfied. If the regularization parameter α is chosen by*

$$\alpha = \frac{\delta^2}{E^2}\left[\psi_p^{-1}\left(\frac{\beta_2\delta^2}{m^2E^2}\right)\right]^{(s-p)/p}, \tag{5.7}$$

then, with the constant $k = \max\{\gamma_1, 1/\sqrt{\beta_2}\}$, we have for $0 \leqslant s < p$ the estimate

$$\left\|x_\alpha^\delta - x^\dagger\right\| \leqslant (k+1)E\sqrt{\psi_p^{-1}\left(\frac{\beta_2\delta^2}{m^2E^2}\right)}. \tag{5.8}$$

Proof. Let us consider the noise amplification error $\|x_\alpha^\delta - x_\alpha\|$ where x_α and x_α^δ are as in (3.1). We use (3.4) and obtain due to A3 (i) the estimates

$$\left\|Ax_\alpha^\delta - Ax_\alpha\right\| \leqslant \delta \qquad \text{and} \qquad \left\|x_\alpha^\delta - x_\alpha\right\|_s \leqslant \gamma_1\delta/\sqrt{\alpha}. \tag{5.9}$$

In case $s > 0$ we follow the proof of proposition 2.1 with $p := s$ and obtain from (5.9)

$$\begin{aligned}\left\|x_\alpha^\delta - x_\alpha\right\| &\leqslant \sup\{\|x\| \mid \|Ax\| \leqslant \delta \wedge \|x\|_s \leqslant \gamma_1\delta/\sqrt{\alpha}\}\\ &\leqslant \frac{k\delta}{\sqrt{\alpha}} \sup\{\|x\| \mid \|Ax\| \leqslant \sqrt{\beta_2\alpha} \wedge \|x\|_s \leqslant 1\}\\ &\leqslant \frac{k\delta}{\sqrt{\alpha}}\sqrt{\psi_s^{-1}\left(\frac{\beta_2\alpha}{m^2}\right)}. \end{aligned} \tag{5.10}$$

Hence, by triangle inequality, the second estimate of (5.9) in case $s = 0$, (5.10) in case $0 < s < p$ and (5.2) we obtain

$$\left\|x_\alpha^\delta - x^\dagger\right\| \leqslant \begin{cases} \dfrac{k\delta}{\sqrt{\alpha}} + E\left[\phi^{-1}\left(\dfrac{\beta_2\alpha}{m^2}\right)\right]^{p/2} & \text{for} \quad s = 0 \\ \dfrac{k\delta}{\sqrt{\alpha}}\sqrt{\psi_s^{-1}\left(\dfrac{\beta_2\alpha}{m^2}\right)} + E\left[\psi_s^{-1}\left(\dfrac{\beta_2\alpha}{m^2}\right)\right]^{p/(2s)} & \text{for} \quad 0 < s < p. \end{cases} \tag{5.11}$$

In the two cases $s = 0$ and $0 < s < p$, the parameter choice (5.7) can be rewritten as

$$\frac{\delta}{\sqrt{\alpha}} = E\left[\phi^{-1}\left(\frac{\beta_2\alpha}{m^2}\right)\right]^{p/2} \qquad \text{and} \qquad \frac{\delta}{\sqrt{\alpha}}\sqrt{\psi_s^{-1}\left(\frac{\beta_2\alpha}{m^2}\right)} = E\left[\psi_s^{-1}\left(\frac{\beta_2\alpha}{m^2}\right)\right]^{p/(2s)}, \tag{5.12}$$

respectively. In the case $s = 0$ we use the first form of (5.12) and obtain from the first part of (5.11) the estimate $\left\|x_\alpha^\delta - x^\dagger\right\| \leqslant (k+1)\delta/\sqrt{\alpha}$. Then, substituting the parameter α of (5.7) with $s = 0$ provides the order optimal error bound (5.8). In the case $0 < s < p$ we use proposition 4.2 and the second form of (5.12), proceed in an analogous way as in case $s = 0$ and obtain again the estimate (5.8). □

5.2. Discrepancy principle

In this subsection we provide order optimal error bounds for methods (3.1) in the low-order case with α chosen *a posteriori* by Morozov's discrepancy principle (3.7). We start by providing some error bound for the total error with respect to the $\|\cdot\|_s$-norm.

Proposition 5.3. *Let x_α^δ be the regularized approximation (3.1) with s chosen such that $0 \leqslant s < p$ and let assumptions A1, A2, A4 and A6 be satisfied. If the regularization parameter α is chosen according to the discrepancy principle (3.7) with $C \geqslant 1$, then*

$$\left\|x_\alpha^\delta - x^\dagger\right\|_s \leqslant E\left[\psi_p^{-1}\left(\frac{(C+1)^2\delta^2}{m^2E^2}\right)\right]^{(p-s)/(2p)}. \tag{5.13}$$

Proof. For α chosen by the discrepancy principle (3.7) the estimate (3.8) attains the form $C^2\delta^2 + \alpha\left\|x_\alpha^\delta - x^\dagger\right\|_s^2 \leqslant \delta^2 + \alpha\left\|R_\alpha^{1/2}B^s x^\dagger\right\|^2$ with $R_\alpha = I - T^*Tg_\alpha(T^*T)$. Since $C \geqslant 1$ we have $\left\|x_\alpha^\delta - x^\dagger\right\|_s^2 \leqslant \left\|R_\alpha^{1/2}B^s x^\dagger\right\|^2$. We exploit in addition the representation $B^s(x^\dagger - x_\alpha) = R_\alpha B^s x^\dagger$, see (3.4), and obtain with assumption A1 the estimate

$$\left\|x_\alpha^\delta - x^\dagger\right\|_s^2 \leqslant \left\|R_\alpha^{1/2}B^s x^\dagger\right\|^2 = \langle B^{2s-p}(x^\dagger - x_\alpha), B^p x^\dagger\rangle \leqslant E\|x_\alpha - x^\dagger\|_{2s-p}. \tag{5.14}$$

Now let us estimate $\|x_\alpha - x^\dagger\|_{2s-p}$ in terms of $\|Ax_\alpha - Ax^\dagger\|$. We multiply (5.6) by $\|z_\alpha\|_{2s-p}/E$ and see that (5.6) may be written in the equivalent form

$$h\left(\frac{\|x_\alpha - x^\dagger\|_{2s-p}}{E}\right) \leqslant \frac{\|Ax_\alpha - Ax^\dagger\|^2}{m^2E^2}$$

with $h(\lambda) = \lambda f(\lambda) = \lambda^{p/(p-s)}\phi(\lambda^{1/(p-s)}) = \psi_p(\lambda^{p/(p-s)})$. The inverse h^{-1} has the representation $h^{-1}(\lambda) = \left[\psi_p^{-1}(\lambda)\right]^{(p-s)/p}$; consequently,

$$\|x_\alpha - x^\dagger\|_{2s-p} \leqslant E\left[\psi_p^{-1}\left(\frac{\|Ax_\alpha - Ax^\dagger\|^2}{m^2E^2}\right)\right]^{(p-s)/p}. \tag{5.15}$$

Let $\widehat{R}_\alpha = I - g_\alpha(TT^*)TT^*$. Then we obtain by exploiting (3.1), (3.4), the triangle inequality, the identity $\widehat{R}_\alpha y^\delta = y^\delta - Ax_\alpha^\delta$ and $|1-\lambda g_\alpha(\lambda)| \leqslant 1$ that follows from A4 that

$$\|Ax_\alpha - Ax^\dagger\| = \|\widehat{R}_\alpha y\| \leqslant \|\widehat{R}_\alpha y^\delta\| + \|\widehat{R}_\alpha(y - y^\delta)\| \leqslant (C+1)\delta. \tag{5.16}$$

Now (5.13) follows from (5.14), (5.15), (5.16) and the monotonicity of ψ_p^{-1}. □

Proposition 5.3 is helpful to derive order optimal error bounds for $\left\|x_\alpha^\delta - x^\dagger\right\|$ in the low-order case provided the regularization parameter has been chosen *a posteriori* by Morozov's discrepancy principle.

Theorem 5.4. *Let x_α^δ be the regularized approximation from (3.1), let $0 \leqslant s < p$ and let assumptions A1, A2, A4 and A6 be satisfied. If the regularization parameter α is chosen according to the discrepancy principle (3.7) with $C \geqslant 1$, then*

$$\left\|x_\alpha^\delta - x^\dagger\right\| \leqslant E\sqrt{\psi_p^{-1}\left(\frac{(C+1)^2\delta^2}{m^2E^2}\right)}. \tag{5.17}$$

Proof. For $s = 0$ the result of theorem 5.4 follows from proposition 5.2. Let $0 < s < p$, let α be chosen according to (3.7) and let $\delta_1 := (C+1)\delta/(mE)$. Then, exploiting (5.13) and the estimate $\left\|Ax_\alpha^\delta - Ax^\dagger\right\| \leqslant \left\|Ax_\alpha^\delta - y^\delta\right\| + \|y - y^\delta\| \leqslant (C+1)\delta$ we have due to proposition 2.1 with $p := s$ that

$$\begin{aligned}\left\|x_\alpha^\delta - x^\dagger\right\| &\leqslant \sup\left\{\|x\| \mid \|Ax\| \leqslant (C+1)\delta \wedge \|x\|_s \leqslant E\left[\psi_p^{-1}\left(\delta_1^2\right)\right]^{(p-s)/(2p)}\right\}\\ &\leqslant E\left[\psi_p^{-1}\left(\delta_1^2\right)\right]^{(p-s)/(2p)}\sqrt{\psi_s^{-1}\left(\delta_1^2\cdot\left[\psi_p^{-1}\left(\delta_1^2\right)\right]^{(s-p)/p}\right)}.\end{aligned}$$

From this estimate and proposition 4.2 we obtain (5.17). □

6. Deconvolution

In this section we discuss a possible application of our results to the deconvolution problem arising in geoscience in the context of models with a non-Wiener filter design (see, e.g. [6, 19]). For example, a standard Gauss–Markov model of satellite observations in the formulation of Bayesian statistics may be written as

$$Az^\dagger = y^\delta + \xi, \tag{6.1}$$

where $z^\dagger$ is the unknown gravity potential which should be recovered from observations y^δ, ξ is a random noise with zero expectation $\mathrm{E}\xi = 0$ and covariance operator $\mathrm{cov}\,\xi = \delta^2 P$. Due to the huge number of observations and unknowns it is reasonable to consider (6.1) as an operator equation in Hilbert spaces with the design operator A acting compactly from the solution space X into the observation space Y. In this context the covariance P can be seen as a bounded self-adjoint nonnegative operator from Y to Y such that for any $g_1, g_2 \in Y$ there holds $\mathrm{E}\langle g_1, \xi\rangle\langle g_2, \xi\rangle = \delta^2\langle Pg_1, g_2\rangle$. If

$$A = \sum_{i=1}^{\infty} a_i v_i \langle u_i, \cdot\rangle$$

is the singular-value decomposition of the design operator, then it is natural to assume that for random noise ξ the Fourier coefficients $\langle v_i, \xi\rangle$ are independent random variables. This assumption allows us to treat the covariance P as a diagonal operator with respect to the system $\{v_i\}$, since for $i \neq j$, $\langle Pv_i, v_j\rangle = \delta^{-2}\mathrm{E}\langle v_i, \xi\rangle\langle v_j, \xi\rangle = \delta^{-2}\mathrm{E}\langle v_i, \xi\rangle\mathrm{E}\langle v_j, \xi\rangle = 0$. Thus,

$$P = \sum_{i=1}^{\infty} p_i v_i \langle v_i, \cdot\rangle,$$

where $p_i = \delta^{-2}\mathrm{E}|\langle v_i, \xi\rangle|^2$. In agreement with the Bayesian approach not only the covariance P is introduced as prior information, but also the expectation $z_0 = \mathrm{E}z^\dagger$, which gives one more observation equation

$$z^\dagger = z_0 + \varepsilon, \qquad \mathrm{E}\varepsilon = 0, \qquad \mathrm{cov}\,\varepsilon = \sigma^2 Q, \qquad Q \in \mathcal{L}(X, X), \qquad Q = Q^* \geqslant 0.$$

Keeping in mind that $z^\dagger \in \overline{R(A^*)}$ it is natural to assume that $\varepsilon = \sum_{i=1}^{\infty} \varepsilon_i u_i$ with independent random Fourier coefficients $\varepsilon_i = \langle u_i, \varepsilon\rangle$. Therefore, as in the case of $\mathrm{cov}\,\xi$,

$$Q = \sum_{i=1}^{\infty} q_i u_i \langle u_i, \cdot\rangle, \qquad q_i = \sigma^{-2}\mathrm{E}|\langle u_i, \varepsilon\rangle|^2. \tag{6.2}$$

Within the framework of Bayesian approach, the estimate $\widehat{z}$ of the unknown element $z^\dagger$ follows from the normal equation

$$(\delta^{-2}A^*P^{-1}A + \sigma^{-2}Q^{-1})\widehat{z} = \delta^{-2}A^*P^{-1}y^\delta + \sigma^{-2}Q^{-1}z_0. \tag{6.3}$$

By introducing $\alpha = \delta^2/\sigma^2$, $x_\alpha^\delta = \widehat{z} - z_0$, $P_* = \sum_{i=1}^{\infty} p_i u_i \langle u_i, \cdot\rangle$ and $B^{2s} = Q^{-1}P_*$, we can reduce equation (6.3) to

$$\alpha B^{2s} x_\alpha^\delta + A^*Ax_\alpha^\delta = A^*(y^\delta - Az_0), \tag{6.4}$$

which is nothing but the Tikhonov regularization, see example 3.1, applied to the equation $Ax^\dagger = y^\delta - Az_0 + \xi$. It allows us to interpret the regularization parameter α as the ratio of the observation noise level δ^2 to the unknown variance σ^2 ([1]). Moreover, in view of the relation $Q = P_*B^{-2s}$ the choice of the prior covariance Q means the choice of the regularizing operator B^{2s}, that is, the choice of the penalty norm in the Tikhonov functional (1.3). For example, self-adaptive regularization (SAR) suggested in [19] is nothing but the Tikhonov regularization corresponding to the choice of prior covariance $Q = A^*P^{-1}A$.

Let the prior covariance Q in (6.2) be chosen in such a way that for $s > 0$ and some continuous strictly monotonically increasing function $\phi : \mathbf{R}^+ \to \mathbf{R}^+$

$$m\sqrt{\phi((q_k/p_k)^{1/s})} \leqslant a_k \leqslant M\sqrt{\phi((q_k/p_k)^{1/s})}, \qquad k = 1, 2, \ldots. \tag{6.5}$$

Then the operator

$$B = (Q^{-1}P_*)^{1/2s} = \sum_{i=1}^{\infty} (p_i/q_i)^{1/2s} u_i \langle u_i, \cdot\rangle$$

automatically meets (4.1). Assume that the covariance P has a finite trace, i.e.

$$\mathrm{E}\|\xi\|^2 = \delta^2 \sum_{i=1}^{\infty} p_i = O(\delta^2).$$

Then it is natural to assume that the norm of actual realization of the random variable $\xi = Az^\dagger - y^\delta = Ax^\dagger + Az_0 - y^\delta$ can be estimated as

$$\|\xi\| = \|Az^\dagger - y^\delta\| = \|Ax^\dagger + Az_0 - y^\delta\| \leqslant c\delta,$$

where c is some fixed constant. In this case from theorems 4.5 and 5.4 we have

Theorem 6.1. *Assume that* $z^\dagger - z_0 \in R((Q^{-1}P_*)^{-p/2s})$ *for some* $p > 0$*. If the function* ϕ *from (6.5) meets assumption A5 (ii) for* $s \geqslant p$*, or assumption A6 for* $s < p$*, then for* $\widehat{z} = z_0 + x_\alpha^\delta$ *and* α *chosen by rule (3.7) with* $y^\delta - Az_0$ *instead of* y^δ *we have*

$$\|z^\dagger - \widehat{z}\| = O\left(\sqrt{\psi_p^{-1}(\delta^2)}\right).$$

Comments. First of all, we would like to note that the reduction of (6.3) to (6.4) was made only for the purposes of analysis. For any σ the estimate $\widehat{z} = \widehat{z}_\sigma = z_0 + x^\delta_{\delta^2/\sigma^2}$ can be found directly from (6.3), and knowledge of the singular-value decompositions of the operators A, P, Q is not necessary for it. The discrepancy principle for $\widehat{z}_\sigma$ consists in choosing the smallest σ such that

$$\|A\widehat{z}_\sigma - y^\delta\| \leqslant C\delta.$$

We would like to stress that this principle does not require any knowledge of p and ϕ. Nevertheless, under the assumptions of theorem 6.1 it automatically provides the best possible order of accuracy. This circumstance seems to be important, because in practice two covariances P and Q reflect our *a priori* knowledge and can be chosen in such a way that the function ϕ from (6.5) related them with design operator A will be rather complicated, but it has not an effect on the computational cost of regularization.

7. Concluding remarks

Regularization of ill-posed operator equations in Hilbert scales is usually studied under the assumption that the operator A involved in the equation and the operator B generating the Hilbert scale are related by some operator-valued index function ϕ. In the classical paper [16] of Natterer, such a relation that characterizes the smoothing properties of A relative to the operator B^{-1} has been expressed in terms of power functions (see (1.4)). Extensions to general index functions have been considered in Mair's paper [9] for the case of high-order regularization in Tikhonov's method. In our paper we have extended Mair's results to a general regularization scheme and to the case of low-order regularization.

Another accomplishment of this paper is the justification of Morozov's discrepancy principle in the light of general index functions ϕ. It is important to note that the discrepancy principle requires neither any knowledge of the index function ϕ nor any knowledge of the solution smoothness measured against the Hilbert scale. Nevertheless, it automatically provides an order optimal choice of the regularization parameter.

Acknowledgments

Parts of this joint work have been done during a stay of the third author at the Johann Radon Institute for Computational and Applied Mathematics, Austrian Academy of Sciences, Linz, Austria, in February 2005. Thanks are due to Professor Heinz W Engl and to Professor Sergei V Pereverzev for the kind invitation and for the hospitality during the visit.

References

[1] Arsenin V Y and Krianev A V 1992 Generalized maximum likelihood method and its application for solving ill-posed problems *Ill-Posed Problems in Natural Sciences* (Moscow: TVP) pp 3–12

[2] Böttcher A, Hofmann B, Tautenhahn U and Yamamoto M 2005 Convergence rates for Tikhonov regularization from different kinds of smoothness conditions *Faculty of Mathematics Univ. Chemnitz Preprint* 2005-9

[3] Engl H W, Hanke M and Neubauer A 1996 *Regularization of Inverse Problems* (Dordrecht: Kluwer)

[4] Hofmann B 1986 *Regularization of Applied Inverse and Ill-Posed Problems* (Leipzig: Teubner)
[5] Hofmann B and Yamamoto M 2005 Convergence rates for Tikhonov regularization based on range inclusions *Inverse Problems* **21** 805–20
[6] Koch K R and Kusche J 2002 Regularization of geopotential determination from satellite data by variance components *J. Geod.* **76** 259–68
[7] Liu F and Nashed M Z 1997 Tikhonov regularization of nonlinear ill-posed poblems with closed operators in Hilbert scales *J. Inverse Ill-Posed Problems* **5** 363–76
[8] Louis A K 1989 *Inverse und schlecht gestellte Probleme* (Stuttgart: Teubner)
[9] Mair B A 1994 Tikhonov regularization for finitely and infinitely smoothing operators *SIAM J. Math. Anal.* **25** 135–47
[10] Mathé P and Pereverzev S V 2003 Geometry of linear ill-posed problems in variable Hilbert scales *Inverse Problems* **19** 789–803
[11] Micchelli C A and Rivlin T J 1977 A survey of optimal recovery *Proc. Int. Symp. Optimal Estimation in Approximation Theory (Freudenstadt, 1976)* (New York: Plenum) pp 1–54
[12] Morozov V A 1966 On the solution of functional equations by the method of regularization *Sov. Math. Dokl.* **7** 414–7
[13] Nair M T 1999 On Morozov's method for Tikhonov regularization as an optimal order yielding algorithm *J. Anal. Appl.* **18** 37–46
[14] Nair M T 2002 Optimal order results for a class of regularization methods using unbounded operators *Integral Eqns Operator Theory* **44** 79–92
[15] Nair M T, Schock E and Tautenhahn U 2003 Morozov's discrepancy principle under general source conditions *J. Anal. Anw.* **1** 199–214
[16] Natterer F 1984 Error bounds for Tikhonov regularization in Hilbert scales *Appl. Anal.* **18** 29–37
[17] Neubauer A 1988 An *a posteriori* parameter choice for Tikhonov regularization in Hilbert scales leading to optimal convergence rates *SIAM J. Numer. Anal.* **25** 1313–26
[18] Neubauer A 1992 Tikhonov regularization of nonlinear ill-posed problems in Hilbert scales *Appl. Anal.* **46** 59–72
[19] Ryzhikov G A and Biryulina M S 1998 Sharp deconvolution in elimination of multiples *68th Ann. Int. Mtg. Soc. Exp. Geophys.* pp 1983–86 (Expanded Abstracts)
[20] Tautenhahn U 1996 Error estimates for regularization methods in Hilbert scales *SIAM J. Numer. Anal.* **33** 2120–30
[21] Tautenhahn U 1998 On a general regularization scheme for nonlinear ill-posed problems: II. Regularization in Hilbert scales *Inverse Problems* **14** 1607–16
[22] Tikhonov A N and Arsenin V Y 1977 *Solution of Ill-Posed Problems* (New York: Wiley)
[23] Vainikko G M and Veretennikov A Y 1986 *Iteration Procedures in Ill-Posed Problems* (Moscow: Nauka) (in Russian)