www.oeaw.ac.at

Johann Radon Institute for
Computational and Applied Mathematics
Austrian Academy of Sciences (ÖAW)

On the generalized discrepancy principle for Tikhonov regularization in Hilbert scales

S. Lu, S. Pereverzyev, Y. Shao, U. Tautenhahn

RICAM-Report 2009-19

www.ricam.oeaw.ac.at

ON THE GENERALIZED DISCREPANCY PRINCIPLE FOR TIKHONOV REGULARIZATION IN HILBERT SCALES

S. LU, S. V. PEREVERZEV, Y. SHAO, AND U. TAUTENHAHN

Dedicated to Charles W. Groetsch

Abstract. For solving linear ill-posed problems regularization methods are required when the right hand side and the operator are with some noise. In the present paper regularized solutions are obtained by Tikhonov regularization in Hilbert scales and the regularization parameter is chosen by the generalized discrepancy principle. Under certain smoothness assumptions we provide order optimal error bounds that characterize the accuracy of the regularized solution. It appears that for getting small error bounds a proper scaling of the penalizing operator B is required. For the computation of the regularization parameter fast algorithms of Newton type are constructed which are based on special transformations. These algorithms are globally and monotonically convergent. The results extend earlier results where the problem operator is exactly given. Some of our theoretical results are illustrated by numerical experiments.

1. Introduction

In this paper we are interested in solving ill-posed problems

$$A_0 x = y_0, \tag{1.1}$$

where $A_0 \in \mathcal{L}(X,Y)$ is a linear, injective and bounded operator with non-closed range $\mathcal{R}(A_0)$ and X, Y are Hilbert spaces with corresponding inner products $(\cdot,\cdot)$ and norms $\|\cdot\|$. Throughout we assume that $y_0 \in \mathcal{R}(A_0)$ so that (1.1) has a unique solution $x^\dagger \in X$. We further assume that (y_0, A_0) are unknown and

(i) $y_\delta \in Y$ is the available noisy right hand side with $\|y_0 - y_\delta\| \le \delta$,
(ii) $A_h \in \mathcal{L}(X,Y)$ is the available noisy operator with $\|A_0 - A_h\| \le h$.

In recent literature, many aspects of treating ill-posed problems with noisy right hand side and noisy operator have been studied, see, e. g., [1, 4, 6, 10, 11, 12, 15, 22, 23, 24, 26, 37, 39, 43, 48]. Ill-posed problems with noisy right hand side and noisy operator arise in different applications. For example, in astronomical observations the point spread function may be changing due to unknown physical conditions leading to a problem with only partially known forward operator. Some special applied ill-posed problems with noisy operators may, e. g., be found in [2, 18, 20, 29, 36].

Date: September 15, 2009.

2000 *Mathematics Subject Classification.* 47A52, 65F22, 65J20, 65M30.

Key words and phrases. Ill-posed problems, inverse problems, noisy right hand side, noisy operator, Tikhonov regularization, Hilbert scales, generalized discrepancy principle, order optimal error bounds, Newton's method, global convergence, monotone convergence.

The numerical treatment of ill-posed problems (1.1) with noisy data (y_δ, A_h) requires the application of special regularization methods. In the method of *Tikhonov regularization in Hilbert scales* a regularized solution $x_\alpha^{\delta,h}$ is obtained by solving the minimization problem

$$\min_{x\in X} J_\alpha(x)\,, \quad J_\alpha(x) = \|A_h x - y_\delta\|^2 + \alpha\|B^s x\|^2, \tag{1.2}$$

where $\alpha > 0$ is the regularization parameter, $B : \mathcal{D}(B) \subset X \to X$ is some unbounded densely defined self-adjoint strictly positive definite operator and s is some generally nonnegative real number that controls the strength of smoothness to be introduced into the regularization method. In many practical problems the operator B is chosen to be a differential operator.

In the special case $h = 0$, Tikhonov regularization in Hilbert scales has been introduced by Natterer [33]. In Natterer's paper it is shown that under the assumptions $\|B^{-a}x\| \sim \|A_0 x\|$ and $\|B^p x^\dagger\| \le E$ the Tikhonov regularized solution $x_\alpha^{\delta,0}$ of the problem (1.2) guarantees order optimal error bounds $\|x_\alpha^{\delta,0} - x^\dagger\| = O(\delta^{p/(a+p)})$ for the p-range $0 < p \le 2s + a$ in case α is chosen *a priori* by $\alpha \sim \delta^{2(a+s)/(a+p)}$. In the meantime regularization in Hilbert scales became quite popular, see, e. g., [34, 38, 40, 41], where method (1.2) has been studied with α chosen *a posteriori* by the discrepancy principle, [5, 41] where method (1.2) has been generalized to a general regularization scheme, [14, 25, 27, 28, 32], where extensions to the case of general source conditions including infinitely smoothing operators A_0 have been treated or [5, 17, 21, 35, 38, 42], where extensions to the nonlinear case may be found. To the authors best knowledge, however, there seem to be no results in the more general case $h \ne 0$.

The accuracy of the regularized solution $x_\alpha^{\delta,h}$ depends on the choice of the regularization parameter. One of the most prominent *a posteriori* rules for choosing α in case of noisy right hand side and noisy operator is the

Generalized discrepancy principle (GDP): Choose $\alpha = \alpha_D$ as the solution of the nonlinear equation

$$\|A_h x_\alpha^{\delta,h} - y_\delta\| = \delta + h\|B^s x_\alpha^{\delta,h}\|. \tag{1.3}$$

This *a posteriori* rule for choosing α goes back to Goncharsky et al. [7, 8]. For $B = I$, the generalized discrepancy principle has intensively been studied by Vainikko in the influential contributions [46, 47, 48]. For the more general case $B \ne I$ some results may be found in [16, 31, 43, 44, 45, 49].

The paper is organized as follows. In Section 2 we give order optimality results for regularized solutions obtained by method (1.2) with α chosen by the generalized discrepancy principle(1.3). In particular, we point out that a proper scaling of the operator B is required and discuss in some detail the standard case $s = 0$. In Section 3 we discuss computational aspects for method (1.2) with the parameter choice (1.3) in the special case $h = 0$. We study properties of equation (1.3) and transform this equation into an equivalent equation with two free parameters (μ, ν). We search for parameters $(\mu, \nu) \subset \mathbb{R}^2$ for which Newton's method for computing the regularization parameter converges globally and monotonically. In Section 4 we extend our results of Section 3 to the more general case $h > 0$ and construct globally convergent Newton type methods for solving the nonlinear equation (1.3).

In the final Section 5 we provide numerical experiments that illustrate some of our theoretical results.

2. **Order optimal error bounds**

In order to guarantee convergence rates for $\|x_\alpha^{\delta,h} - x^\dagger\|$, certain smoothness assumptions are necessary which we formulate in terms of some densely defined unbounded self-adjoint strictly positive operator $B : X \to X$. We introduce a *Hilbert scale* $(X_r)_{r\in\mathbb{R}}$ induced by B which is the completion of $\cap_{k=0}^{\infty}\mathcal{D}(B^k)$ with respect to the Hilbert space norm

$$\|x\|_r = \|B^r x\|, \quad r \in \mathbb{R}$$

and consider the following two classical assumptions.

Assumption A1. For some positive constants m and a we assume the *link condition*

$$m\|x\|_{-a} \le \|A_0 x\| \quad \text{for all } x \in X.$$

Assumption A2. For some positive constants E and p we assume the *solution smoothness* $x^\dagger = B^{-p}v$ with $v \in X$ and $\|v\| \le E$, that is,

$$x^\dagger \in M_{p,E} = \Big\{x \in X \,\Big|\, \|x\|_p \le E\Big\}.$$

Assumption A1 characterizes the smoothing properties of the operator A_0 relative to the operator B^{-1}, and Assumption A2 characterizes the smoothness of the unknown solution $x^\dagger$ allowing the study of different smoothness situations for $x^\dagger$. It can be shown that under a two-sided link condition $\|A_0 x\| \sim \|x\|_{-a}$ and Assumption A2, the best possible worst case error for identifying $x^\dagger$ from noisy data (y_δ, A_h) is of the order $O\left((\delta + h)^{p/(p+a)}\right)$. From [45] we know that the regularized solution $x_\alpha^{\delta,h}$ with α chosen by the generalized discrepancy principle provides the optimal order for $s = p$. Since p is generally unknown there arises the question about order optimal error bounds if regularization is carried out with $s \ne p$. An order optimality proof for the p-range $p \in [1, 2 + a]$ in case $s = 1$ may be found in [43]. We follow this way of proof, exploit the interpolation inequality

$$\|z\|_r \le \|z\|_{-a}^{(s-r)/(s+a)} \|z\|_s^{(a+r)/(s+a)} \tag{2.1}$$

which holds true for any $r \in [-a, s]$, $a + s \ne 0$ (see, e. g., [19]) and obtain

Theorem 2.1. *Let $\|B^{-1}\| \le 1$, let Assumptions A1 and A2 with $p \in [s, 2s + a]$ be satisfied and let $x_\alpha^{\delta,h}$ be the Tikhonov regularized solution of problem (1.2) with α chosen by the generalized discrepancy principle (1.3). Then,*

$$\|x_\alpha^{\delta,h} - x^\dagger\| \le 2E^{\frac{a}{p+a}} \left(\frac{\delta + h\|x^\dagger\|_s}{m}\right)^{\frac{p}{p+a}}. \tag{2.2}$$

Proof. In our *first step* of the proof we show that for α chosen by (1.3) we have

$$\|x_\alpha^{\delta,h}\|_s \le \|x^\dagger\|_s. \tag{2.3}$$

For the proof of (2.3) we use $J_\alpha(x_\alpha^{\delta,h}) \le J_\alpha(x^\dagger)$ and obtain due to the GDP (1.3), the triangle inequality, $0 \le s \le p$ and $\|B^{-1}\| \le 1$ that

$$\begin{aligned}\left(\delta + h\|x_\alpha^{\delta,h}\|_s\right)^2 + \alpha\|x_\alpha^{\delta,h}\|_s^2 &\le \left(\delta + h\|x^\dagger\|\right)^2 + \alpha\|x^\dagger\|_s^2 \\ &\le \left(\delta + h\|x^\dagger\|_s\right)^2 + \alpha\|x^\dagger\|_s^2.\end{aligned}$$

Since $t \to (\delta + ht)^2 + \alpha t^2$ is increasing we obtain (2.3). In our *second step* of the proof we show that for every element $x \in X$ with $\|x\|_s \le \|x^\dagger\|_s$ we have under the side conditions $p \in [s, 2s+a]$, $a > 0$ and $\|x^\dagger\|_p \le E$ the estimate

$$\|x - x^\dagger\| \le (2E)^{a/(p+a)}\|x - x^\dagger\|_{-a}^{p/(p+a)}. \tag{2.4}$$

For the proof of (2.4) we introduce the abbreviation $z := x^\dagger - x$ and derive three estimates. Due to $\|x\|_s \le \|x^\dagger\|_s$ and Cauchy-Schwarz inequality we have a first estimate

$$\|z\|_s^2 \le 2\left(B^s x^\dagger, B^s z\right) = \left(B^p x^\dagger, B^{2s-p} z\right) \le 2E\|z\|_{2s-p}. \tag{2.5}$$

From (2.1) with $r := 2s - p$ we have a second estimate

$$\|z\|_{2s-p} \le \|z\|_{-a}^{(p-s)/(s+a)}\|z\|_s^{(a+2s-p)/(s+a)}. \tag{2.6}$$

A further application of (2.1) with $r := 0$ gives a third estimate

$$\|z\| \le \|z\|_{-a}^{s/(s+a)}\|z\|_s^{a/(s+a)}. \tag{2.7}$$

Now, a proper combination of the three estimates (2.5)–(2.7) gives (2.4). In our *third step* of the proof we derive an estimate for $\|x_\alpha^{\delta,h} - x^\dagger\|_{-a}$. Due to Assumption A1, the triangle inequality, the GDP (1.3), $\|B^{-1}\| \le 1$ and estimate (2.3) we obtain

$$\begin{aligned}\|x_\alpha^{\delta,h} - x^\dagger\|_{-a} &\le \frac{1}{m}\|A_h(x_\alpha^{\delta,h} - x^\dagger)\| \\ &\le \frac{1}{m}\left(\delta + h\|x_\alpha^{\delta,h}\| + \|A_h x_\alpha^{\delta,h} - y_\delta\|\right) \\ &\le \frac{1}{m}\left(2\delta + 2h\|x^\dagger\|_s\right).\end{aligned} \tag{2.8}$$

Now, estimate (2.2) follows from (2.4) with $x = x_\alpha^{\delta,h}$ and (2.8). □

From Theorem 2.1 we obtain

Corollary 2.2. *Let $x_\alpha^{\delta,h}$ be the Tikhonov regularized solution of problem (1.2) with $s = 0$, let α be chosen by the generalized discrepancy principle (1.3) with $s = 0$ and let $x^\dagger$ obey $x^\dagger = (A^*A)^{p/2}v$ with $\|v\| \le E$. Then, for $p \in (0, 1]$,*

$$\|x_\alpha^{\delta,h} - x^\dagger\| \le 2E^{\frac{1}{p+1}}\left(\delta + h\|x^\dagger\|\right)^{\frac{p}{p+1}}. \tag{2.9}$$

Proof. For the choice $B = (A^*A)^{-1/2}$, Assumption A2 is equivalent to the source condition $x^\dagger = (A^*A)^{p/2}v$ with $\|v\| \le E$ and A1 holds true with $a = 1$ and $m = 1$. Hence, the result of Corollary 2.2 follows from Theorem 2.1. □

Remark 2.3. The order optimality result $\|x_\alpha^{\delta,h} - x^\dagger\| = O\left((\delta+h)^{p/(p+1)}\right)$ of Corollary 2.2 may also be found in [48]. The proof in [48] is done for a general regularization scheme and requires to choose α from the nonlinear equation

$$\|A_h x_\alpha^{\delta,h} - y_\delta\| = C\left(\delta + h\|x_\alpha^{\delta,h}\|\right)$$

with some $C > 1$. The convergence rate proof in [48] is more complicated as our proof and provides compared with our estimate (2.9) larger constants that even depend on h and are therefore only valid for h sufficiently small.

Now we consider without loss of generality the special case $s = 1$ and ask the question if replacing B by βB with some constant β influences the accuracy of the regularized solution. The answer is *yes* in the case $h \neq 0$ for the regularized solution of problem (1.2) with α chosen by the generalized discrepancy principle (1.3). Assume that $x_{\alpha,\beta}^{\delta,h}$ is obtained by solving

$$\min_{x\in X} J_\alpha(x), \quad J_\alpha(x) = \|A_h x - y_\delta\|^2 + \alpha\|\beta B x\|^2 \tag{2.10}$$

with α chosen by the generalized discrepancy principle, that is, $\alpha = \alpha_D$ is the solution of the equation

$$\|A_h x_{\alpha,\beta}^{\delta,h} - y_\delta\| = \delta + h\|\beta B x_{\alpha,\beta}^{\delta,h}\|. \tag{2.11}$$

Then we observe two limit relations:

Proposition 2.4. *Let $x_{\alpha,\beta}^{\delta,h}$ be given by (2.10) with $\alpha = \alpha_D$ chosen by the generalized discrepancy principle (2.11). Then, following two limit relations are valid:*

(i) *For $\beta \to \infty$ we have $x_{\alpha,\beta}^{\delta,h} \to 0$.*

(ii) *For $\beta \to 0$ we have $x_{\alpha,\beta}^{\delta,h} \to x_\gamma^{\delta,h}$ where $x_\gamma^{\delta,h} = (A_h^* A_h + \gamma B^* B)^{-1} A_h^* y_\delta$ and γ is the solution of the equation $\|A_h x_\gamma^{\delta,h} - y_\delta\| = \delta$.*

The observation in Proposition 2.4 has consequences. A wrong choice of β leads to a bad regularized solution $x_{\alpha,\beta}^{\delta,h}$. For β chosen too large, the regularized solution is close to zero, whereas for β chosen too small, the regularized solution is generally highly oscillating. As a result, there exists an optimal β-value for which the total error becomes minimal. The error bound in Theorem 2.1 tells us that $\beta = 1/\|B^{-1}\|$ seems to be a good *a priori* choice.

3. Tikhonov regularization in the special case $h = 0$

In this section we discuss computational aspects for method (1.2) with the parameter choice (1.3) in the special case $h = 0$. Without loss of generality we restrict our considerations to the special case $s = 1$. In this special case, the regularized solution of problem (1.2) with A_h replaced by A_0 will be denoted by x_α^δ. For computing this regularized solution with $\alpha = \alpha_D$ chosen by the discrepancy principle (1.3), we observe that $\alpha = \alpha_D$ may be found by solving the nonlinear equation

$$f(\alpha) := \|A_0 x_\alpha^\delta - y_\delta\|^2 - \delta^2 = 0. \tag{3.1}$$

Our next proposition tells us that $f : \mathbb{R}^+ \to \mathbb{R}$ is monotonically increasing and that equation (3.1) possesses a unique positive solution $\alpha_D > 0$ provided

$$\|Py_\delta\| < \delta < \|y_\delta\|. \tag{3.2}$$

Here P is the orthogonal projector onto $\mathcal{R}(T)^\perp$ and T is given by $T = A_0 B^{-1}$.

Proposition 3.1. *Let $x_\alpha^\delta = (A_0^* A_0 + \alpha B^* B)^{-1} A_0^* y_\delta$, let f be defined by (3.1) and let $v_\alpha^\delta = (A_0^* A_0 + \alpha B^* B)^{-1} B^* B x_\alpha^\delta$. Then:*

(i) *$f : \mathbb{R}^+ \to \mathbb{R}$ is continuous and obeys the limit relations*

$$\lim_{\alpha\to 0} f(\alpha) = \|Py_\delta\|^2 - \delta^2 \quad \textit{and} \quad \lim_{\alpha\to\infty} f(\alpha) = \|y_\delta\|^2 - \delta^2.$$

(ii) *$f : \mathbb{R}^+ \to \mathbb{R}$ is monotonically increasing and its derivative is given by*

$$f'(\alpha) = 2\alpha (Bv_\alpha^\delta, Bx_\alpha^\delta) > 0. \tag{3.3}$$

(iii) *$f : \mathbb{R}^+ \to \mathbb{R}$ is convex for small α-values, but concave for large α-values. Its second derivative is given by*

$$f''(\alpha) = 2(Bv_\alpha^\delta, Bx_\alpha^\delta) - 6\alpha (Bv_\alpha^\delta, Bv_\alpha^\delta). \tag{3.4}$$

(iv) *Assume that the data y_δ obey (3.2). Then the equation $f(\alpha) = 0$ possesses a unique positive solution $\alpha_D > 0$.*

The proof of Proposition 3.1 is standard and may be derived from results in [5]. From property (iii) we conclude that global and monotone convergence of Newton's method for solving equation (3.1) cannot be guaranteed. In the literature, different alternatives for solving nonlinear equations of the type (3.1) have been proposed:

(1) In [9], see also [5, Prop. 9.8], the function $g(r) := f(r^{-1})$ is introduced. This function appears to be decreasing and convex. As a consequence, Newton's method for solving $g(r) = 0$ converges for arbitrary positive starting values $r_0 < r_D$ globally and monotonically from the left to the unique solution $r_D = \alpha_D^{-1}$.

(2) In the trust region version of the Gauss-Newton method for solving nonlinear least squares problems, a trust region step requires to solve for given Δ the equation $\|x_\alpha^\delta\| = \Delta$. This can effectively be realized by solving the equivalent *secular* equation $h(\alpha) := \|x_\alpha^\delta\|^{-1} - \Delta^{-1} = 0$ by Newton's method, see [30] and [3, Subsection 7.3.3].

The above two ideas motivate us

(1) to introduce the function $h : \mathbb{R}^+ \to \mathbb{R}$ by $h(\alpha) := \|A_0 x_\alpha^\delta - y_\delta\|^\mu - \delta^\mu$,
(2) to introduce the function $g : \mathbb{R}^+ \to \mathbb{R}$ by $g(r) := h(r^\nu)$,
(3) to consider the nonlinear equation

$$g(r) := h(r^\nu) = \|A_0 x_{r^\nu}^\delta - y_\delta\|^\mu - \delta^\mu = 0 \tag{3.5}$$

and to ask following question: *For which pairs $(\mu, \nu) \subset \mathbb{R}^2$ it can be guaranteed that Newton's method applied to the nonlinear equation $g(r) = 0$ converges globally and monotonically to the unique solution $r_D = \alpha_D^{1/\nu}$ of equation (3.5)?*

To answer this question we start by computing the first two derivatives of g.

Proposition 3.2. *Let $x = x^\delta_{r^\nu}$ be the solution of $(A_0^*A_0 + r^\nu B^*B)x = A_0^* y_\delta$ and $v = v^\delta_{r^\nu}$ be the solution of $(A_0^*A_0 + r^\nu B^*B)v = B^*Bx^\delta_{r^\nu}$. Then the first and second derivative of the function $g : \mathbb{R}^+ \to \mathbb{R}$ defined by (3.5) are given by*

$$g'(r) = \mu\nu r^{2\nu-1}(Bv, Bx)\|A_0x - y_\delta\|^{\mu-2} \tag{3.6}$$

and

$$\begin{aligned} g''(r) &= \mu(\mu-2)\nu^2 r^{2\nu-2}(A_0v, A_0x - y_\delta)^2\|A_0x - y_\delta\|^{\mu-4} \\ &\quad + \mu\nu(2\nu-1)r^{2\nu-2}\|A_0v\|^2\|A_0x - y_\delta\|^{\mu-2} \\ &\quad - \mu\nu(\nu+1)r^{3\nu-2}\|Bv\|^2\|A_0x - y_\delta\|^{\mu-2}. \end{aligned} \tag{3.7}$$

Proof. The function g possesses the representation

$$g(r) = f_1^{\mu/2}(r^\nu) - \delta^\mu \quad \text{with} \quad f_1(\alpha) = \|A_0x^\delta_\alpha - y_\delta\|^2.$$

For the first derivative we have

$$g'(r) = \frac{\mu}{2}\nu r^{\nu-1} f_1^{\mu/2-1}(r^\nu) f_1'(r^\nu).$$

We use the identity $f_1' = f'$, exploit that f' is given by (3.3) and obtain (3.6). For the second derivative of g we have

$$\begin{aligned} g''(r) = &\frac{\mu}{2}\nu(\nu-1)r^{\nu-2} f_1^{\mu/2-1}(r^\nu) f_1'(r^\nu) \\ &+ \frac{\mu}{2}\left(\frac{\mu}{2} - 1\right)\nu^2 r^{2\nu-2} f_1^{\nu/2-2}(r^\nu) f_1'^2(r^\nu) \\ &+ \frac{\mu}{2}\nu^2 r^{2\nu-2} f_1^{\mu/2-1}(r^\nu) f_1''(r^\nu). \end{aligned}$$

We use the identities $f_1' = f'$ and $f_1'' = f''$, exploit that f' and f'' are given by (3.3) and (3.4), respectively, and obtain

$$\begin{aligned} g''(r) &= \mu\nu(\nu-1)r^{2\nu-2}\|A_0x - y_\delta\|^{\mu-2}(Bv, Bx) \\ &\quad + \mu(\mu-2)\nu^2 r^{4\nu-2}\|A_0x - y_\delta\|^{\mu-4}(Bv, Bx)^2 \\ &\quad + \mu\nu^2 r^{2\nu-2}\|A_0x - y_\delta\|^{\mu-2}\left((Bv, Bx) - 3r^\nu\|Bv\|^2\right) \\ &= \mu(\mu-2)\nu^2 r^{4\nu-2}\|A_0x - y_\delta\|^{\mu-4}(Bv, Bx)^2 \\ &\quad + \mu\nu(2\nu-1)r^{2\nu-2}\|A_0x - y_\delta\|^{\mu-2}(Bv, Bx) \\ &\quad - 3\mu\nu^2 r^{3\nu-2}\|A_0x - y_\delta\|^{\mu-2}\|Bv\|^2. \end{aligned}$$

We rewrite the first summand by using the identity $(Bv, Bx) = r^{-\nu}(A_0v, y_\delta - A_0x)$, rewrite the second summand by using the identity $(Bv, Bx) = \|A_0v\|^2 + r^\nu\|Bv\|^2$, collect terms and obtain (3.7). $\square$

The use of formulas (3.6) and (3.7) allows us to search for (μ, ν)-domains $G \subset \mathbb{R}^2$ with non-changing sign for the derivatives g' and g''. In particular, we will show that the situation of Figure 1 is valid. In the proof which is given in the next proposition we exploit in some parts of $G = \cup_{i=1}^4 G_i$ that due to Cauchy-Schwarz inequality we have

$$(A_0v, A_0x - y_\delta) \le \|A_0v\|\|A_0x - y_\delta\|. \tag{3.8}$$

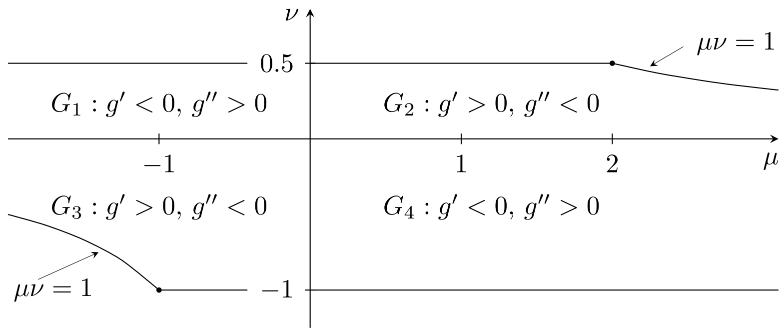

Figure 1: (μ, ν) – domain G with non-changing sign for the derivatives g' and g''

Proposition 3.3. *Let* G_1 – G_4 *be the domains of Figure 1. Then,* $g : \mathbb{R}^+ \to \mathbb{R}$ *defined by (3.5) obeys*

(i) $g' < 0$ *and* $g'' > 0$ *for* $(\mu, \nu) \in G_1 \cup G_4$ *and*
(ii) $g' > 0$ *and* $g'' < 0$ *for* $(\mu, \nu) \in G_2 \cup G_3$.

Proof. For the first and second derivative of g we use the formulas (3.6) and (3.7) of Proposition 3.2, respectively, observe that $(Bv, Bx) > 0$, decompose the second derivative into the sum $g''(r) = s_1 + s_2 + s_3$ and distinguish four cases.

Case $(\mu, \nu) \in G_1 = \{(\mu, \nu) \in \mathbb{R}^2 \,|\, -\infty < \mu < 0 \,\wedge\, 0 < \nu \le \frac{1}{2}\}$: In this case we have $g' < 0$, $s_1 > 0$, $s_2 \ge 0$ and $s_3 > 0$, which proves part (i) for $(\mu, \nu) \in G_1$.

Case $(\mu, \nu) \in G_2 = \{(\mu, \nu) \in \mathbb{R}^2 \,|\, 0 < \mu < \infty \,\wedge\, 0 < \nu \le \frac{1}{2} \,\wedge\, \mu\nu \le 1\}$: In this case we have $g' > 0$, $s_1 < 0$ for $\mu < 2$, $s_1 \ge 0$ for $\mu \ge 2$, $s_2 \le 0$ and $s_3 < 0$. Hence, in the subcase $\mu < 2$ we have $g''(r) < 0$. In the subcase $\mu \ge 2$ we use (3.8) and obtain $s_1 \le \mu(\mu - 2)\nu^2 r^{2\nu-2} \|A_0 v\|^2 \|A_0 x - y_\delta\|^{\mu-2}$. Consequently,

$$s_1 + s_2 \le \mu\nu(\mu\nu - 1) r^{2\nu-2} \|A_0 v\|^2 \|A_0 x - y_\delta\|^{\mu-2} \le 0, \tag{3.9}$$

which yields $g''(r) < 0$ and proves part (ii) for $(\mu, \nu) \in G_2$.

Case $(\mu, \nu) \in G_3 = \{(\mu, \nu) \in \mathbb{R}^2 \,|\, -\infty < \mu < 0 \,\wedge\, -1 \le \nu < 0 \,\wedge\, \mu\nu \ge 1\}$: In this case we have $g' > 0$, $s_1 > 0$, $s_2 < 0$ and $s_3 \le 0$. Due to (3.8), the first summand can be estimated by

$$s_1 \le \mu(\mu - 2)\nu^2 r^{2\nu-2} \|A_0 v\|^2 \|A_0 x - y_\delta\|^{\mu-2},$$

which yields (3.9). Hence, $g''(r) < 0$, which proves part (ii) for $(\mu, \nu) \in G_3$.

Case $(\mu, \nu) \in G_4 = \{(\mu, \nu) \in \mathbb{R}^2 \,|\, 0 < \mu < \infty \,\wedge\, -1 \le \nu < 0\}$: In this case we have $g' < 0$, $s_1 \ge 0$ for $\mu \ge 2$, $s_1 < 0$ for $\mu < 2$, $s_2 > 0$ and $s_3 \ge 0$. Hence, in the subcase $\mu \ge 2$ we have $g''(r) > 0$. In the subcase $\mu < 2$ we use (3.8) and obtain $s_1 \ge \mu(\mu - 2)\nu^2 r^{2\nu-2} \|A_0 v\|^2 \|A_0 x - y_\delta\|^{\mu-2}$. Consequently,

$$s_1 + s_2 \ge \mu\nu(\mu\nu - 1) r^{2\nu-2} \|A_0 v\|^2 \|A_0 x - y_\delta\|^{\mu-2} > 0,$$

which yields $g''(r) > 0$ and proves part (i) for $(\mu, \nu) \in G_4$. $\square$

In the next proposition we formulate conditions under which Newton's method for solving nonlinear equations converges globally and monotonically.

Proposition 3.4. *Let $g : \mathbb{R}^+ \to \mathbb{R}$ be twice continuously differentiable and assume that the equation $g(r) = 0$ has a unique solution $r_D > 0$. Assume further that the starting value r_0 obeys $0 < r_0 < r_D$ and that either*

$$\text{(i) } g' < 0 \text{ and } g'' > 0 \qquad \text{or} \qquad \text{(ii) } g' > 0 \text{ and } g'' < 0.$$

Then, Newton's method for solving $g(r) = 0$ converges globally and monotonically from the left and the speed of convergence is locally quadratic.

Due to formula (3.6), Newton's method $r_{k+1} = r_k - g(r_k)/g'(r_k)$, $k = 0, 1, 2, \ldots$, for solving the nonlinear equation (3.5) possesses the form

$$r_{k+1} = r_k - \frac{\|A_0 x^\delta_{r_k^\nu} - y_\delta\|^\mu - \delta^\mu}{\mu\nu r_k^{2\nu-1}(Bv^\delta_{r_k^\nu}, Bx^\delta_{r_k^\nu})\|A_0 x^\delta_{r_k^\nu} - y_\delta\|^{\mu-2}}. \tag{3.10}$$

From Propositions 3.3 and 3.4 we obtain that this iteration method converges monotonically from the left for arbitrary starting values $r_0 \in (0, r_D)$ and arbitrary $(\mu, \nu) \in G = \cup_{i=1}^4 G_i$, which is the main result of this section.

Theorem 3.5. *Let α_D be the solution of equation (3.1), $r_D := \alpha_D^{1/\nu}$ be the solution of equation (3.5), $(\mu, \nu) \in \cup_{i=1}^4 G_i$ and G_1 – G_4 the domains of Figure 1. Then, Newton's method (3.10) for solving equation (3.5) converges globally and monotonically from the left for starting values $0 < r_0 < r_D$. In particular,*

(1) *for $(\mu, \nu) \in G_1 \cup G_2$ and $0 < \alpha_0 < \alpha_D$, the sequence $(\alpha_k) := (r_k^\nu)$ converges monotonically from the left to α_D,*
(2) *for $(\mu, \nu) \in G_3 \cup G_4$ and $\alpha_0 > \alpha_D$, the sequence $(\alpha_k) := (r_k^\nu)$ converges monotonically from the right to α_D.*

Remark 3.6. We made numerical experiments to check for which (μ, ν) the Newton iteration (3.10) gives fast convergence of the sequence $(\alpha_k) := (r_k^\nu)$. We found that in the domain $(\mu, \nu) \in G_1 \cup G_2$ fast convergence is guaranteed for $(\mu, \nu) = (2, 0.5)$ and that in the domain $(\mu, \nu) \in G_3 \cup G_4$ fast convergence is guaranteed for $(\mu, \nu) = (-1, -1)$. Due to this observation and the results of Theorem 3.5 we propose following strategy of applying Newton's method (3.10) where we have global convergence for arbitrary starting values $\alpha_0 > 0$:

(i) Choose $\alpha_0 > 0$ and compute the discrepancy $d = \|A_0 x^\delta_{\alpha_0} - y_\delta\|$. Then, depending on the magnitude of d, we proceed either according to (ii) or according to (iii).
(ii) If $d < \delta$, then we know from Proposition 3.1 that $\alpha_0 < \alpha_D$. In this case, Theorem 3.5 tells us that for $(\mu, \nu) \in G_1 \cup G_2$ the sequence $(\alpha_k) := (r_k^\nu)$ converges monotonically from the left to α_D. Hence, in case $d < \delta$ we start the Newton iteration (3.10) with $(\mu, \nu) = (2, 0.5)$.
(iii) If $d > \delta$, then we know from Proposition 3.1 that $\alpha_0 > \alpha_D$. In this case, Theorem 3.5 tells us that for $(\mu, \nu) \in G_3 \cup G_4$ the sequence $(\alpha_k) := (r_k^\nu)$ converges monotonically from the right to α_D. Hence, in case $d > \delta$ we start the Newton iteration (3.10) with $(\mu, \nu) = (-1, -1)$.

For $s = 1$ in equation (1.3), the results of Theorem 3.5 and Remark 3.6 lead us to following algorithm.

Algorithm 1 Global convergent Newton iteration for equation (1.3) with $h = 0$.

Input: $\varepsilon > 0$, y_δ, A_0, B, δ and $\alpha > 0$.

1: Solve $(A_0^* A_0 + \alpha B^* B)x = A_0^* y_\delta$ and compute $d := \|A_0 x - y_\delta\|$.

2: **if** $d < \delta$ **then** $\mu := 2$, $\nu := \frac{1}{2}$, $r := \alpha^{1/\nu}$ **else** $\mu := -1$, $\nu := -1$, $r := \alpha^{1/\nu}$.

3: Solve $(A_0^* A_0 + \alpha B^* B)v = B^* Bx$ and compute $s := (v, B^* Bx)$.

4: Update $r_{\text{new}} := r - \dfrac{d^\mu - \delta^\mu}{\mu\nu r^{2\nu-1} s d^{\mu-2}}$.

5: **if** $|r_{\text{new}} - r| \geq \varepsilon |r|$ **then**
$r := r_{\text{new}}$, $\alpha := r^\nu$, $x := (A_0^* A_0 + \alpha B^* B)^{-1} A_0^* y_\delta$, $d := \|A_0 x - y_\delta\|$
and **goto** 3 **else** stop.

4. Tikhonov regularization in the general case $h \neq 0$

In this section we discuss computational aspects for the method (1.2) with the parameter choice (1.3) in the general case $h \neq 0$. Again, without loss of generality, we restrict our considerations to the case $s = 1$. For properties of equation (1.3) and conditions under which this equation possesses a unique solution α_D we consider the equivalent equation

$$f(\alpha) = \|A_h x_\alpha^{\delta,h} - y_\delta\|^2 - \left(\delta + h\|Bx_\alpha^{\delta,h}\|\right)^2 = 0. \tag{4.1}$$

Our next proposition tells us that f is monotonically increasing and that equation (4.1) possesses a unique positive solution $\alpha_D > 0$ provided

$$\|P_h y_\delta\| - h\|x_{\delta,h}^\dagger\| < \delta < \|y_\delta\|. \tag{4.2}$$

Here P_h is the orthogonal projector onto $\mathcal{R}(T_h)^\perp$, T_h is given by $T_h = A_h B^{-1}$ and $x_{\delta,h}^\dagger$ is the Moore-Penrose solution of the perturbed linear system $T_h x = y_\delta$ (if it exists). If $x_{\delta,h}^\dagger$ does not exists, then $\|Bx_\alpha^{\delta,h}\| \to \infty$ for $\alpha \to 0$ and the left inequality of (4.2) is automatically satisfied.

Proposition 4.1. *Let f be defined by (4.1), let $x_\alpha^{\delta,h}$ be the solution of (1.2) with $s = 1$, and let $v_\alpha^{\delta,h} = (A_h^* A_h + \alpha B^* B)^{-1} B^* B x_\alpha^{\delta,h}$. Then:*

(i) *$f : \mathbb{R}^+ \to \mathbb{R}$ is continuous and obeys the limit relations*

$$\lim_{\alpha\to 0} f(\alpha) = \|P_h y_\delta\|^2 - \left(\delta + h\|x_{\delta,h}^\dagger\|\right)^2 \quad \textit{and} \quad \lim_{\alpha\to\infty} f(\alpha) = \|y_\delta\|^2 - \delta^2.$$

(ii) *$f : \mathbb{R}^+ \to \mathbb{R}$ is monotonically increasing and its derivative is given by*

$$f'(\alpha) = 2\left(\alpha + h^2 + h\delta/\|Bx_\alpha^{\delta,h}\|\right)(Bv_\alpha^{\delta,h}, Bx_\alpha^{\delta,h}) > 0.$$

(iii) *Assume that (4.2) holds. Then the equation $f(\alpha) = 0$ possesses a unique positive solution $\alpha_D > 0$.*

The proof of Proposition 4.1 is analogous to [24, Prop. 4.5], where the special case $B = I$ has been treated. Now, analogously to Section 3 we introduce the function $h : \mathbb{R}^+ \to \mathbb{R}$ by

$$h(\alpha) := \|A_h x_\alpha^{\delta,h} - y_\delta\|^\mu - (\delta + h\|Bx_\alpha^{\delta,h}\|)^\mu$$

where $x_\alpha^{\delta,h}$ is the solution of the operator equation $(A_h^* A_h + \alpha B^* B)x = A_h^* y_\delta$, transform (1.3) into an equivalent equation

$$g(r) := h(r^\nu) = \|A_h x_{r^\nu}^{\delta,h} - y_\delta\|^\mu - \left(\delta + h\|B x_{r^\nu}^{\delta,h}\|\right)^\mu = 0 \tag{4.3}$$

with two free parameters (μ,ν) and ask, as in Section 3, the following question: *For which pairs $(\mu,\nu) \subset \mathbb{R}^2$ it can be guaranteed that Newton's method applied to the nonlinear equation $g(r) = 0$ converges globally and monotonically to the unique solution $r_D = \alpha_D^{1/\nu}$ of equation (4.3)?*

To answer this question, we decompose the functions h and g into the sum $h = h_1 + h_2$ and $g = g_1 + g_2$, respectively, where

$$\begin{aligned} g_1(r) &= h_1(r^\nu) = \|A_h x_{r^\nu}^{\delta,h} - y_\delta\|^\mu, \\ g_2(r) &= h_2(r^\nu) = -\left(\delta + h\|B x_{r^\nu}^{\delta,h}\|\right)^\mu. \end{aligned} \tag{4.4}$$

We observe that for the derivatives of g_1 there hold analogous formulas as given in Proposition 3.2. For the first two derivatives of the function g_2 we have

Proposition 4.2. *Let $x = x_{r^\nu}^{\delta,h}$ be the solution of $(A_h^* A_h + r^\nu B^* B)x = A_h^* y_\delta$ and $v = v_{r^\nu}^{\delta,h}$ be the solution of $(A_h^* A_h + r^\nu B^* B)v = B^* B x_{r^\nu}^{\delta,h}$. Then the first and second derivative of the function $g_2 : \mathbb{R}^+ \to \mathbb{R}$ defined by (4.4) are given as follows:*

$$g_2'(r) = h\mu\nu r^{\nu-1} \left(\delta + h\|Bx\|\right)^{\mu-1} \|Bx\|^{-1}(Bv, Bx) \tag{4.5}$$

and

$$\begin{aligned} g_2''(r) = c^2 \Big[&\mu\nu^2 (Bv, Bx)^2 \left(\delta + h\|Bx\|\right) - h\mu(\mu-1)\nu^2 \|Bx\|(Bv, Bx)^2 \\ &- 3\mu\nu^2 \|Bx\|^2 \|Bv\|^2 \left(\delta + h\|Bx\|\right) \\ &+ \mu\nu(\nu-1) r^{-\nu} \|Bx\|^2 (Bv, Bx) \left(\delta + h\|Bx\|\right) \Big] \end{aligned} \tag{4.6}$$

with $c^2 = hr^{2\nu-2}\|Bx\|^{-3}(\delta + h\|Bx\|)^{\mu-2}$.

Proof. Consider the equation $(A_h^* A_h + \alpha B^* B)x_\alpha^{\delta,h} = A_h^* y_\delta$. Differentiating both sides by α yields

$$B^* B x_\alpha^{\delta,h} + (A_h^* A_h + \alpha B^* B)\frac{d}{d\alpha} x_\alpha^{\delta,h} = 0,$$

or equivalently,

$$\frac{d}{d\alpha} x_\alpha^{\delta,h} = -(A_h^* A_h + \alpha B^* B)^{-1} B^* B x_\alpha^{\delta,h} =: -v_\alpha^{\delta,h}.$$

Consequently,

$$\frac{d}{d\alpha}\|B x_\alpha^{\delta,h}\| = \frac{d}{d\alpha}\left(\|B x_\alpha^{\delta,h}\|^2\right)^{1/2} = -\|B x_\alpha^{\delta,h}\|^{-1}(B v_\alpha^{\delta,h}, B x_\alpha^{\delta,h}). \tag{4.7}$$

Consider the equation $(A_h^* A_h + \alpha B^* B)v_\alpha^{\delta,h} = B^* B x_\alpha^{\delta,h}$. Differentiating both sides by α yields

$$B^* B v_\alpha^{\delta,h} + (A_h^* A_h + \alpha B^* B)\frac{d}{d\alpha} v_\alpha^{\delta,h} = B^* B \frac{d}{d\alpha} x_\alpha^{\delta,h},$$

or equivalently,

$$\begin{aligned}\frac{d}{d\alpha}v_\alpha^{\delta,h} &= (A_h^*A_h+\alpha B^*B)^{-1}\Big(B^*B\frac{d}{d\alpha}x_\alpha^{\delta,h} - B^*Bv_\alpha^{\delta,h}\Big)\\ &= -2(A_h^*A_h+\alpha B^*B)^{-1}B^*Bv_\alpha^{\delta,h}.\end{aligned}$$

Consequently,

$$\begin{aligned}\frac{d}{d\alpha}(Bv_\alpha^{\delta,h}, Bx_\alpha^{\delta,h}) &= \Big(\frac{d}{d\alpha}v_\alpha^{\delta,h}, B^*Bx_\alpha^{\delta,h}\Big) + \Big(B^*Bv_\alpha^{\delta,h}, \frac{d}{d\alpha}x_\alpha^{\delta,h}\Big)\\ &= -\,3\|Bv_\alpha^{\delta,h}\|^2. \end{aligned} \tag{4.8}$$

Now we introduce the function $f_2(\alpha) = \delta + h\|Bx_\alpha^{\delta,h}\|$. Due to (4.7), the first derivative is given by

$$f_2'(\alpha) = -h\frac{(Bv_\alpha^{\delta,h}, Bx_\alpha^{\delta,h})}{\|Bx_\alpha^{\delta,h}\|}. \tag{4.9}$$

From (4.7), (4.8), (4.9) and quotient rule we obtain

$$f_2''(\alpha) = -h\frac{(Bv_\alpha^{\delta,h}, Bx_\alpha^{\delta,h})^2 - 3\|Bv_\alpha^{\delta,h}\|^2\|Bx_\alpha^{\delta,h}\|^2}{\|Bx_\alpha^{\delta,h}\|^3}. \tag{4.10}$$

The functions g_2 and f_2 are related by $g_2(r) = -f^\mu(r^\nu)$. Consequently,

$$g_2'(r) = -\mu\nu r^{\nu-1} f_2^{\mu-1}(r^\nu) f_2'(r^\nu). \tag{4.11}$$

Substituting f_2 and (4.9) into (4.11) gives (4.5). From (4.11) we have

$$\begin{aligned}g_2''(r) = &-\mu\nu(\nu-1)r^{\nu-2}f_2^{\mu-1}(r^\nu)f_2'(r^\nu)\\ &-\mu(\mu-1)\nu^2r^{2\nu-2}f_2^{\mu-2}(r^\nu)f_2'^2(r^\nu)\\ &-\mu\nu^2r^{2\nu-2}f_2^{\mu-1}(r^\nu)f_2''(r^\nu).\end{aligned} \tag{4.12}$$

Substituting f_2, (4.9) and (4.10) into (4.12) gives (4.6). □

The use of formulas (4.5) and (4.6) allows us to search for (μ,ν)-domains $H \subset \mathbb{R}^2$ with non-changing sign for the derivatives g_2' and g_2''. In particular, we will show that the situation of Figure 2 is valid. In the proof which is given in the next proposition we exploit in some parts of $H = \cup_{i=1}^4 H_i$ that by Cauchy-Schwarz inequality we have

$$(Bv, Bx) \le \|Bv\|\|Bx\|. \tag{4.13}$$

Proposition 4.3. *Let H_1 – H_4 be the domains of Figure 2. Then, $g_2 : \mathbb{R}^+ \to \mathbb{R}$ defined by (4.4) obeys*

(i) $g_2' < 0$ *and* $g_2'' > 0$ *for* $(\mu,\nu) \in H_1 \cup H_4$ *and*
(ii) $g_2' > 0$ *and* $g_2'' < 0$ *for* $(\mu,\nu) \in H_2 \cup H_3$.

Proof. Scalar multiplication of the equation $(A_h^*A_h + r^\nu B^*B)v = B^*Bx$ by v yields $(Bv, Bx) = \|A_hv\|^2 + r^\nu\|Bv\|^2$. We substitute this expression into the third summand of (4.6), collect terms and obtain

$$g''(r) = c^2\,(\delta E + h\|Bx\|F)$$

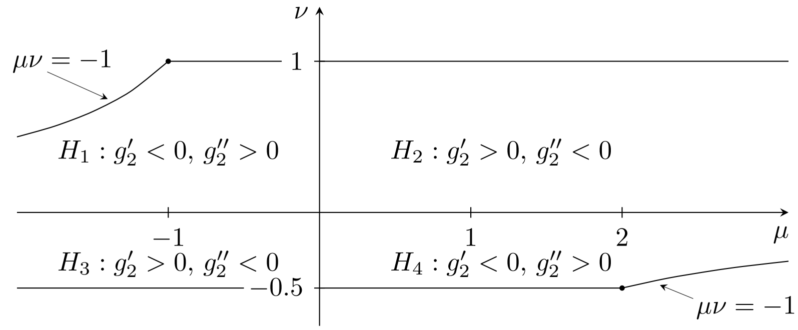

Figure 2: (μ, ν) – domain H with non-changing sign for the derivatives g_2' and g_2''

where c^2 is given in Proposition 4.2 and

$$\begin{aligned} E = {} & \mu\nu^2(Bv, Bx)^2 - \mu\nu(2\nu+1)\|Bv\|^2\|Bx\|^2 \\ & + \mu\nu(\nu-1)r^{-\nu}\|A_h v\|^2\|Bx\|^2, \end{aligned}$$

$$\begin{aligned} F = {} & -\mu(\mu-2)\nu^2(Bv, Bx)^2 - \mu\nu(2\nu+1)\|Bv\|^2\|Bx\|^2 \\ & + \mu\nu(\nu-1)r^{-\nu}\|A_h v\|^2\|Bx\|^2. \end{aligned}$$

We write both expressions E and F in the form

$$E = s_1 + s_2 + s_3, \quad F = s_4 + s_5 + s_6,$$

use for the first derivative of g_2 the formula (4.5) and distinguish four cases.

Case $(\mu, \nu) \in H_1 = \{(\mu, \nu) \in \mathbb{R}^2 \mid -\infty < \mu < 0 \wedge 0 < \nu \leq 1 \wedge \mu\nu + 1 \geq 0\}$: In this case we have $g_2' < 0$, $s_1 < 0$, $s_2 > 0$ and $s_3 \geq 0$. Due to (4.13), s_1 can be estimated by $s_1 \geq \mu\nu^2\|Bv\|^2\|Bx\|^2$. Hence,

$$s_1 + s_2 \geq -\mu\nu(\nu+1)\|Bv\|^2\|Bx\|^2 > 0,$$

which implies $E > 0$. Furthermore, $s_4 < 0, s_5 > 0$ and $s_6 \geq 0$. Due to (4.13), $s_4 \geq -\mu(\mu-2)\nu^2\|Bv\|^2\|Bx\|^2$. Hence,

$$s_4 + s_5 \geq -\mu\nu(\mu\nu+1)\|Bv\|^2\|Bx\|^2 \geq 0, \tag{4.14}$$

which gives $F \geq 0$ and proves part (i) for $(\mu, \nu) \in H_1$.

Case $(\mu, \nu) \in H_2 = \{(\mu, \nu) \in \mathbb{R}^2 \mid 0 < \mu < \infty \wedge 0 < \nu \leq 1\}$: In this case we have $g_2' > 0$, $s_1 > 0$, $s_2 < 0$ and $s_3 \leq 0$. We use (4.13) and obtain $s_1 \leq \mu\nu^2\|Bv\|^2\|Bx\|^2$. Consequently,

$$s_1 + s_2 \leq -\mu\nu(\nu+1)\|Bv\|^2\|Bx\|^2 < 0,$$

which yields $E < 0$. Furthermore, we have $s_4 < 0$ for $\mu > 2$, $s_4 \geq 0$ for $\mu \leq 2$, $s_5 < 0$ and $s_6 \leq 0$. Hence, in the subcase $\mu > 2$ we have $F < 0$. In the subcase $\mu \leq 2$ we estimate s_4 by $s_4 \leq -\mu(\mu-2)\nu^2\|Bv\|^2\|Bx\|^2$ and obtain

$$s_4 + s_5 \leq -\mu\nu(\mu\nu+1)\|Bv\|^2\|Bx\|^2 < 0,$$

which gives $F < 0$ and proves part (ii) for $(\mu, \nu) \in H_2$.

Case $(\mu,\nu) \in H_3 = \{(\mu,\nu) \in \mathbb{R}^2 \,|\, -\infty < \mu < 0 \wedge -\frac{1}{2} \le \nu < 0\}$: In this case we have $g_2' > 0$, $s_1 < 0$, $s_2 \le 0$ and $s_3 < 0$, which gives $E < 0$. Furthermore, we have $s_4 < 0$, $s_5 \le 0$ and $s_6 < 0$, which gives $F < 0$ and proves part (ii) for $(\mu,\nu) \in H_3$.

Case $(\mu,\nu) \in H_4 = \{(\mu,\nu) \in \mathbb{R}^2 \,|\, 0 < \mu < \infty \wedge -\frac{1}{2} \le \nu < 0 \wedge \mu\nu + 1 \ge 0\}$: In this case we have $g_2' < 0$, $s_1 > 0$, $s_2 \ge 0$ and $s_3 > 0$, which yields $E > 0$. Furthermore, $s_4 > 0$ for $\mu < 2$, $s_4 \le 0$ for $\mu \ge 2$, $s_5 \ge 0$ and $s_6 > 0$. Hence, in the subcase $\mu < 2$ we have $F > 0$. In the subcase $\mu \ge 2$ we use (3.8) and obtain

$$s_4 \ge -\mu(\mu-2)\nu^2 \|Bv\|^2 \|Bx\|^2.$$

From this estimate we obtain (4.14). This estimate yields $F > 0$ and proves part (i) for $(\mu,\nu) \in H_4$. □

Due to formulae (3.6) and (4.5), Newton's method $r_{k+1} = r_k - g(r_k)/g'(r_k)$, $k = 0,1,2,\dots$, for solving the nonlinear equation (4.1) possesses the form

$$r_{k+1} = r_k - \frac{\|A_h x - y_\delta\|^\mu - (\delta + h\|Bx\|)^\mu}{\mu\nu r_k^{\nu-1}(Bv,Bx)\left(r_k^\nu \|A_h x - y_\delta\|^{\mu-2} + h\|Bx\|^{-1}(\delta + h\|Bx\|)^{\mu-1}\right)}$$

with $x := x^{\delta,h}_{r_k^\nu}$ and $v := v^{\delta,h}_{r_k^\nu}$. From Propositions 3.3, 3.4 and 4.3 we obtain that this iteration method converges monotonically from the left for arbitrary starting values $r_0 \in (0, r_D)$ and arbitrary $(\mu,\nu) \in G \cap H$, where G is given in Figure 1 and H is given in Figure 2.

Theorem 4.4. *Let α_D be the solution of equation (1.3), $r_D := \alpha_D^{1/\nu}$ be the solution of equation (4.3) and $(\mu,\nu) \in G \cap H$ where G and H are the domains of Figure 1 and Figure 2. Then, Newton's method for solving equation (4.3) converges globally and monotonically from the left for starting values $r_0 < r_D$. In particular,*

(1) *for $(\mu,\nu) \in (G_1 \cup G_2) \cap (H_1 \cup H_2)$ and $\alpha_0 < \alpha_D$, the sequence $(\alpha_k) := (r_k^\nu)$ converges monotonically from the left to α_D,*
(2) *for $(\mu,\nu) \in (G_3 \cup G_4) \cap (H_3 \cup H_4)$ and $\alpha_0 > \alpha_D$, the sequence $(\alpha_k) := (r_k^\nu)$ converges monotonically from the right to α_D.*

Remark 4.5. We made numerical experiments, see Section 5, to check for which (μ,ν) the Newton iteration for solving equation (4.3) gives fast convergence of the sequence $(\alpha_k) := (r_k^\nu)$. We found that in the domain $(\mu,\nu) \in (G_1 \cup G_2) \cap (H_1 \cup H_2)$ fast convergence is guaranteed for $(\mu,\nu) = (2, 0.5)$ and that in the domain $(\mu,\nu) \in (G_3 \cup G_4) \cap (H_3 \cup H_4)$ fast convergence is guaranteed for $(\mu,\nu) = (-2,-0.5)$. This observation and the results of Theorem 4.4 lead us, as outlined in Remark 3.6, to Algorithm 2 for solving equation (1.3) with $s = 1$. This algorithm converges globally and monotonically for arbitrary starting values $\alpha_0 > 0$.

Algorithm 2 Global convergent Newton iteration for solving equation (1.3).

Input: $\varepsilon > 0$, y_δ, A_h, B, δ, h and $\alpha > 0$.

1: Solve $(A_h^* A_h + \alpha B^* B)x = A_h^* y_\delta$ and compute $d := \|A_h x - y_\delta\|$, $n := \|Bx\|$.

2: **if** $d < \delta + hn$ **then** $\mu := 2$, $\nu := \frac{1}{2}$, $r := \alpha^{1/\nu}$
else $\mu := -2$, $\nu := -\frac{1}{2}$, $r := \alpha^{1/\nu}$.

3: Solve $(A_h^* A_h + \alpha B^* B)v = B^* Bx$ and compute $s := (Bv, Bx)$, $n := \|Bx\|$.

4: Update $r_{\text{new}} := r - \dfrac{d^\mu - (\delta + hn)^\mu}{\mu\nu r^{\nu-1} s\Big(r^\nu d^{\mu-2} + hn^{-1}(\delta + hn)^{\mu-1}\Big)}$.

5: **if** $|r_{\text{new}} - r| \geq \varepsilon |r|$ **then**
$r := r_{\text{new}}$, $\alpha := r^\nu$, $x := (A_h^* A_h + \alpha B^* B)^{-1} A_h^* y_\delta$, $d := \|A_h x - y_\delta\|$
and **goto** 3 **else** stop.

5. Numerical experiments

In this section we provide different numerical experiments. In the first two subsections we provide our test examples and discuss how we choose B. In a third subsection we perform experiments that confirm the facts mentioned in the Remark 4.5. In a fourth subsection we illustrate the theoretical results of the order optimal error bounds of Theorem 2.1 and in a fifth subsection we investigate the infuence of a second parameter β as discussed in Proposition 2.4.

5.1. Test examples.

As test examples we use approximations of the first kind Fredholm integral equation

$$[\mathcal{A}x](s) := \int_0^1 K(s,t)x(t)\,\mathrm{d}t = y(s), \quad 0 \leq s \leq 1, \tag{5.1}$$

$\mathcal{A} : L^2(0,1) \to L^2(0,1)$, leading to ill-conditioned linear systems of equations. Introducing the nodes $t_j = s_j = j\tau$, $j = 0, \ldots, n$, with step size $\tau = 1/n$, and searching for discretized solutions $x(t) = \sum_{j=1}^n x_j \varphi_j(t)$ with zero order spline basis functions

$$\varphi_j(t) = \begin{cases} 1/\sqrt{\tau} & \text{for } t \in [t_{j-1}, t_j] \\ 0 & \text{for } t \notin [t_{j-1}, t_j] \end{cases}$$

leads to the Galerkin approximation $A_0 x = y$ for (5.1) with $A_0 = (a_{ij})$,

$$a_{ij} = \langle \mathcal{A}\varphi_j, \varphi_i \rangle = \int_0^1 \int_0^1 K(s,t)\varphi_i(s)\varphi_j(t)\,\mathrm{d}s\,\mathrm{d}t \approx \tau K\Big(s_i - \frac{\tau}{2}, t_j - \frac{\tau}{2}\Big), \tag{5.2}$$

$x = (x_j)$, $y = (y_i)$ and $y_i = \langle y(s), \varphi_i(s) \rangle = \int_0^1 y(s)\varphi_i(s)\,\mathrm{d}s \approx \sqrt{\tau}\, y\,(s_i - \tau/2)$.

Example 5.1. In our first test example we use for A_0 the matrix with elements (5.2), for $x^\dagger$ the vector with coordinates $x_j := \sqrt{\tau} x(t_j - \tau/2)$ and for y_0 the vector $y_0 := A_0 x^\dagger$. For the functions in (5.1) we use

$$K(s,t) = \begin{cases} s(1-t) & \text{for } s \leq t \\ t(1-s) & \text{for } s \geq t, \end{cases} \qquad x(t) = 4t(1-t), \quad y(s) = \frac{s}{3}(s^3 - 2s^2 + 1).$$

The matrix $-A_0$ can be generated by the Matlab function *deriv*2 from [13].

Example 5.2. Our second test example is analogous to Example 5.1, however, instead of $x(t)$ and $y(s)$ we use

$$x(t) = t \quad \text{and} \quad y(s) = \frac{s}{6}(s^2 - 1).$$

We note that by the finite dimensional approximations in Examples 5.1 and 5.2 it is guaranteed that

(i) $\|A_0\|_F \approx \|\mathcal{A}\|_{HS} = \sqrt{\int_0^1 \int_0^1 K^2(s,t)\,\mathrm{d}s\,\mathrm{d}t}$ holds and that

(ii) $\|x_0\|_2 \approx \|x(t)\|_{L^2(0,1)}$ and $\|y_0\|_2 \approx \|y(s)\|_{L^2(0,1)}$ holds.

For modeling noise in the right hand side y_0 and in the matrix A_0, for given nonnegative σ_y and σ_A we compute

$$y_\delta = y_0 + \sigma_y \frac{\|y_0\|_2}{\|e\|_2}\, e, \qquad \text{and} \qquad A_h = A_0 + \sigma_A \frac{\|A_0\|_F}{\|E\|_F}\, E,$$

where $e = (e_i)$ is a random vector with $e_i \sim \mathcal{N}(0,1)$ and $E = (e_{ij})$ is a random matrix with $e_{ij} \sim \mathcal{N}(0,1)$. In this way of modeling noise we guarantee that for the relative errors we have $\|y_0 - y_\delta\|_2/\|y_0\|_2 = \sigma_y$ and $\|A_0 - A_h\|_F/\|A_0\|_F = \sigma_A$. The noise levels δ and h are then given by

$$\delta = \sigma_y\, \|y_0\|_2 \quad \text{and} \quad h = \sigma_A\, \|A_0\|_F.$$

For $\sigma_y = 0.03$ the vectors $\sqrt{n}\cdot y_0$ and $\sqrt{n}\cdot y_\delta$ are displayed in Figure 3 and for $\sigma_A = 0.03$ the matrices $n \cdot A_0$ and $n \cdot A_h$ are displayed in Figure 4.

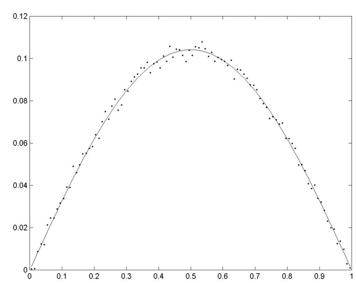

Figure 3: Exact and noisy right hand side for Example 5.1, $\sigma_y = 0.03$, $n = 100$

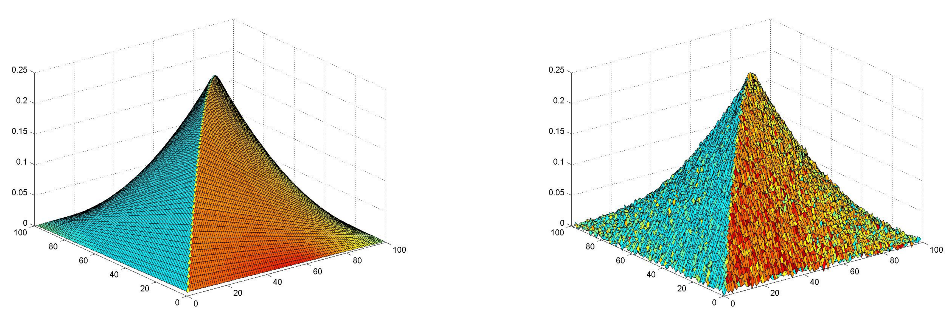

Figure 4: Exact and noisy matrix (left/right) for Example 5.1, $\sigma_A = 0.03$, $n = 100$

In Figure 5 we display the exact solution $\sqrt{n}\cdot x^\dagger$ and different regularized solutions $\sqrt{n}\cdot x_\alpha^{\delta,h}$ for Example 5.1 with $\sigma_y = 0.03$, $\sigma_A = 0.03$, $B = I$ and $n = 100$. In this example we have $\delta \approx 0.00222$ and $h \approx 0.00316$. It is easy to see that $x^\dagger$

can well be approximated by $x_\alpha^{\delta,h}$ with properly chosen α, and that $x_\alpha^{\delta,h}$ is highly oscillating for small α, while for large α the regularized solution is close to zero.

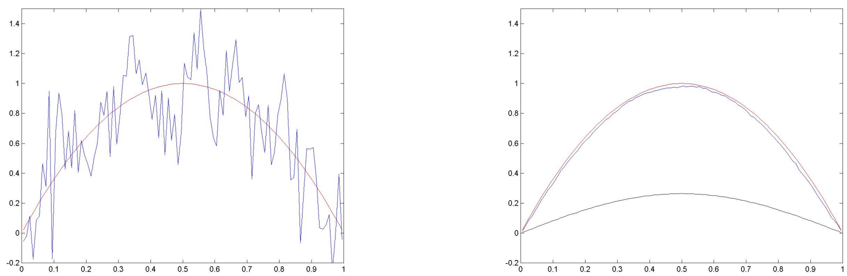

Figure 5: Exact solution $x^\dagger$ and regularized solutions $x_\alpha^{\delta,h}$ for Example 5.1 with $B = I$, $\sigma_y = 0.03$, $\sigma_A = 0.03$ and $n = 100$. Left: $x^\dagger$ and $x_\alpha^{\delta,h}$ with $\alpha = 0.000003$. Right: $x^\dagger$ and $x_\alpha^{\delta,h}$ with $\alpha = 0.0003$ and $\alpha = 0.03$

5.2. **Choosing the operator** B**.** For $B : D \subset L^2(0,1) \to L^2(0,1)$ we choose

$$Bx = \sum_{k=1}^{\infty} k(x, e_k)e_k \quad \text{with} \quad e_k(t) = \sqrt{2}\sin(k\pi t). \tag{5.3}$$

Checking Asumptions A1 and A2 we have

Proposition 5.3. *Let $B : D \subset L^2(0,1) \to L^2(0,1)$ be defined by (5.3), then:*

(i) *The operator $\mathcal{A}$ defined by (5.1) with the kernel function of Example 5.1 obeys Assumption A1 with $m = \pi^{-2}$ and $a = 2$.*
(ii) *The function $x(t) = 4t(1-t)$ of Example 5.1 obeys A2 for all $p \in [0, \frac{5}{2})$.*
(iii) *The function $x(t) = t$ of Example 5.2 obeys Assumption A2 for all $p \in [0, \frac{1}{2})$.*

We note that the operator $B^2 : D \subset L^2(0,1) \to L^2(0,1)$ is the second order differential operator

$$[B^2x](t) := -\pi^{-2}x''(t), \quad D(B) = \left\{x \in H^2(0,1) : x(0) = 0,\, x(1) = 0\right\}.$$

The discrete approximations for B^2 and B are given by the matrices B_2 and B_1, respectively, where

$$B_2 = \begin{pmatrix} 2 & -1 & & \\ -1 & \ddots & \ddots & \\ & \ddots & \ddots & -1 \\ & & -1 & 2 \end{pmatrix} \quad \text{and} \quad B_1 = B_2^{1/2}.$$

For the smallest eigenvalue $\lambda_{\min}$ of B_2 there holds $\lambda_{\min} = 2\left(1 - \cos\frac{\pi}{n+1}\right) \approx \frac{\pi^2}{(n+1)^2}$. Hence, in order to guarantee the assumption $\|B^{-1}\| \le 1$ in Theorem 2.1, we will work in our experiments with $B := \frac{n+1}{\pi}B_1$.

5.3. **Number of iterations.** In this subsection we perform experiments that confirm the facts mentioned in the Remark 4.5. All experiments have been done with $s = 1$. From Theorem 4.4 we know that Newton's method for solving equation (4.1) converges globally for any $(\mu, \nu) \in G \cap H$, where for $\nu > 0$ we have monotone convergence from the left, while for $\nu < 0$ we have monotone convergence from the right with respect to α. We made different experiments and collect two of them in Table 1 and Table 2. From our experiments we found the pair $(\mu, \nu) = (2, \frac{1}{2})$ in the range $\nu > 0$ and the pair $(\mu, \nu) = (-2, -\frac{1}{2})$ in the range $\nu < 0$, which provide the smallest number of iterations compared with other pairs. Due to these numerical results, we have used these two pairs in Algorithm 2.

ν	$\mu = -4.5$	$\mu = -4$	$\mu = -3.5$	$\mu = -3$	$\mu = -2.5$	$\mu = -2$
$-1/\mu$	24	22	20	17	15	13
$-0.5/\mu$	32	29	26	22	19	16
ν	$\mu = -2$	$\mu = -1.2$	$\mu = -0.4$	$\mu = 0.4$	$\mu = 1.2$	$\mu = 2$
0.50	13	11	9	7	6	**5**
0.35	14	12	10	9	7	6
0.20	18	16	13	11	10	9
0.05	37	33	30	27	24	22
ν	$\mu = 2$	$\mu = 2.5$	$\mu = 3$	$\mu = 3.5$	$\mu = 4$	$\mu = 4.5$
$1/\mu$	5	5	6	6	6	7
$0.5/\mu$	8	8	9	10	10	11

Table 1: Iteration numbers in the range $G \cap H$ with $\nu > 0$ for Example 5.1 with $B := \frac{n+1}{\pi} B_1$, $\sigma_y = 0$, $\sigma_A = 0.03$, $n = 200$, $\varepsilon = 0.001$ and $\alpha_0 = \alpha_D/100$

ν	$\mu = -4.5$	$\mu = -4$	$\mu = -3.5$	$\mu = -3$	$\mu = -2.5$	$\mu = -2$
$0.5/\mu$	18	16	15	13	12	10
$1/\mu$	12	11	10	9	8	7
ν	$\mu = -2$	$\mu = -1.2$	$\mu = -0.4$	$\mu = 0.4$	$\mu = 1.2$	$\mu = 2$
-0.05	27	27	29	32	36	39
-0.20	12	12	13	14	17	19
-0.35	9	9	10	11	13	16
-0.50	**7**	7	8	10	12	14
ν	$\mu = 2$	$\mu = 2.5$	$\mu = 3$	$\mu = 3.5$	$\mu = 4$	$\mu = 4.5$
$-0.5/\mu$	18	21	24	28	31	34
$-1/\mu$	14	17	19	22	24	27

Table 2: Iteration numbers in the range $G \cap H$ with $\nu < 0$ for Example 5.1 with $B := \frac{n+1}{\pi} B_1$, $\sigma_y = 0$, $\sigma_A = 0.03$, $n = 200$, $\varepsilon = 0.001$ and $\alpha_0 = 100\alpha_D$

5.4. **Accuracy of the regularized solutions.** In this subsection we illustrate the order optimal error bounds mentioned in the Theorem 2.1. From worst case analysis, Proposition 5.3 and Theorem 2.1 we conlude

(i) For B chosen by (5.3), the best possible error bound for identifying the function $x(t) = 4t(1-t)$ of Example 5.1 from noisy data (y_δ, A_h) is of order $O((\delta + h)^q)$ for any $q < \frac{5}{9}$. Choosing $s = 1$, this rate can be obtained by method (1.2) with the parameter choice (1.3).

(ii) For B chosen by (5.3), the best possible error bound for identifying the function $x(t) = t$ of Example 5.2 from noisy data (y_δ, A_h) is of the order $O((\delta + h)^q)$ for any $q < \frac{1}{5}$. Choosing $s = 1$, the assumption $p \in [1, 2 + a]$ in Theorem 2.1 is violated and we cannot conclude that method (1.2) with the parameter choice (1.3) provides the best possible order. Therefore, we will check this by numerical experiments.

In our numerical experiments the regularization parameter α_D has been computed by Algorithm 2 with $\varepsilon = 0.001$. In order to keep the discretization error small we have used the dimension number $n = 400$ in all computations. We note that for both Examples 5.1 and 5.2 we performed computations with $\sigma_y = 0$ and different σ_A. In all examples, the matrix A_0 has been randomly perturbed 20 times. For every perturbed matrix A_h the regularization parameters α_D and the regularized solutions have been computed, and the error values in Tables 3 and 4 represent corresponding mean values. In Table 3 we added the theoretically error bound $\|x^{0,h}_{\alpha_D} - x^\dagger\|_{L^2(0,1)} \le 8\left(\frac{2}{\pi\sqrt{3}}\right)^{1/2} \cdot \sqrt{h} \approx 1.58\sigma_A^{1/2} := e_{\text{theor}}$ that follows from the error bound of Theorem 2.1 with $p = 2$.

σ_A	α_D	$\alpha_D/h^{4/3}$	e_D	$e_D/h^{5/9}$	e_{theor}
5.0 E−2	5.33 E−4	0.581	.0442	0.814	.3515
1.0 E−2	1.01 E−4	0.938	.0243	1.096	.1572
5.0 E−3	4.37 E−5	1.026	.0205	1.355	.1112
1.0 E−3	4.32 E−6	0.868	.0077	1.242	.0497
5.0 E−4	1.81 E−6	0.913	.0050	1.191	.0352
1.0 E−4	2.34 E−7	1.013	.0019	1.132	.0157
5.0 E−5	9.92 E−8	1.081	.0013	1.117	.0111
1.0 E−5	1.46 E−8	1.358	.0007	1.449	.0050

Table 3: Regularization parameters α_D and errors $e_D := \|x^{0,h}_{\alpha_D} - x^\dagger\|_2$ for Example 5.1 with $B = \frac{n+1}{\pi} B_1$ and $n = 400$

Both Tables 3 and 4 show following:

(i) For the Example 5.1 with α_D chosen by the generalized discrepancy principle (1.3), the error $\|x^{0,h}_{\alpha_D} - x^\dagger\|$ obeys the predicted rate $O(h^{p/(p+a)}) = O(h^{5/9})$ of Theorem 2.1, and α_D tends to zero with the rate $O(h^{2(a+1)/(a+p)}) = O(h^{4/3})$.

(ii) For the Example 5.2 with α_D chosen by the generalized discrepancy principle (1.3), the error $\|x^{0,h}_{\alpha_D} - x^\dagger\|$ obeys the expected rate $O(h^{p/(p+a)}) = O(h^{1/5})$

σ_A	α_D	α_D/h^2	e_D	$e_D/h^{1/5}$
5.0 E−2	9.87 E−05	3.554	.2840	0.811
1.0 E−2	4.71 E−06	4.240	.2198	0.866
5.0 E−3	1.23 E−06	4.429	.1960	0.887
1.0 E−3	5.24 E−08	4.717	.1495	0.933
5.0 E−4	1.34 E−08	4.825	.1327	0.952
1.0 E−4	5.59 E−10	5.032	.1000	0.989
5.0 E−5	1.42 E−10	5.113	.0882	1.002
1.0 E−5	5.90 E−12	5.311	.0649	1.017

Table 4: Regularization parameters α_D and errors $e_D := \|x^{0,h}_{\alpha_D} - x^\dagger\|_2$ for Example 5.2 with $B = \frac{n+1}{\pi} B_1$ and $n = 400$

of Theorem 2.1, and α_D tends to zero not with the rate $O(h^{2(a+1)/(a+p)}) = O(h^{12/5})$, but with the rate $O(h^2)$. However, for this example, the assumption $p \in [1, 2+a]$ of Theorem 2.1 is violated.

5.5. **Proper scaling of B.** In this subsection we show by experiment the influence of replacing B_1 by βB_1 as discussed at the end of Section 2. In different experiments we observed following:

n	α_D	$e(\beta = 2 + \frac{n}{50})$	α_D	$e(\beta = \frac{n+1}{\pi})$
20	7.27 E−4	.0198	2.96 E−4	.0272
50	9.94 E−4	.0145	2.96 E−4	.0275
100	1.03 E−3	.0124	2.97 E−4	.0299
200	1.20 E−3	.0115	2.97 E−4	.0316
400	1.58 E−3	.0117	2.97 E−4	.0326

Table 5: Errors $e(\beta) := \|x^{\delta,h}_{\alpha,\beta} - x^\dagger\|_2$ and regularization parameters α_D for $\beta = 2 + \frac{n}{50}$ (left) and $\beta = \frac{n+1}{\pi}$ (right) for Example 5.1 with $B := \beta B_1$, $\sigma_y = 0$ and $\sigma_A = 0.03$ (mean values in case of 20 random experiments)

(i) There exists an optimal parameter β_{opt} for which $e(\beta) := \|x^{\delta,h}_{\alpha,\beta} - x^\dagger\|_2$ as a function of β becomes minimal.
(ii) Due to the limit relations (i) and (ii) of Proposition 2.4, the error $e(\beta)$ is growing for growing β-values $\beta > \beta_{\text{opt}}$ and also growing for decreasing β-values $\beta < \beta_{\text{opt}}$.
(iii) We observed that for growing dimension numbers n the optimal parameter β_{opt} is growing.
(iv) We do not know how to determine β_{opt}. In Table 5, a statistical experiment with 20 random examples shows that for Example 5.1 the *a priori* parameter choice $\beta := 2 + \frac{n}{50}$ provides better results than the *a priori* parameter

choice $\beta := \frac{n+1}{\pi} \approx 1/\|B_1^{-1}\|$ which obeys the assumption $\|(\beta B_1)^{-1}\| \leq 1$ of Theorem 2.1.

Acknowledgments. This joint work has been conducted during the *Mini Special Semester on Inverse Problems*, May 18th – July 15th, 2009, organized by Johann Radon Institute for Computational and Applied Mathematics (RICAM), Austrian Academy of Sciences. U. Tautenhahn as a long-term guest thanks RICAM, and in particular Prof. Dr. Heinz W. Engl, for kind invitation and hospitality during the visit. The first and second authors are supported by the Austrian Fonds Zur Förderung der Wissenschaftlichen Forschung (FWF), Grant P20235-N18.

References

1. L. Cavalier and N. Hengartner, *Adaptive estimation for inverse problems with noisy operators*, Inverse Problems **21** (2005), 1345–1361.
2. D. Colton, M. Piana, and R. Potthast, *A simple method using Morozov's discrepancy principle for solving inverse scattering problems*, Inverse Problems **13** (1997), 1477–1493.
3. A. R. Conn, N. I. M. Gould, and P. L. Toint, *Trust-Region Methods*, SIAM, Philadelphia, 2000.
4. S. Efromovich and V. Koltchinskii, *On inverse problems with unknown operators*, IEEE Trans. Inform. Theory **47** (2001), 2876–2894.
5. H. W. Engl, M. Hanke, and A. Neubauer, *Regularization of Inverse Problems*, Kluwer, Dordrecht, 1996.
6. G. H. Golub, P. C. Hansen, and D. P. O'Leary, *Tikhonov regularization and total least squares*, SIAM J. Matrix Anal. Appl. **21** (1999), 185–194.
7. A. V. Goncharsky, A. S. Leonov, and A. G. Yagola, *A regularizing algorithm for incorrectly formulated problems with an approximately specified operator*, Zh. Vychisl. Mat. Mat. Fiz. **12** (1972), 1592–1594.
8. ———, *Generalized discrepancy principle*, Zh. Vychisl. Mat. Mat. Fiz. **13** (1973), 294–302.
9. V. I. Gordonova, , and V. A. Morozov, *Numerical parameter selection algorithms in the regularization methods*, Zh. Vychisl. Mat. Mat. Fiz. **13** (1974), 1–9.
10. C. W. Groetsch, *The Theory of Tikhonov Regularization for Fredholm Integral Equations of the First Kind*, Pitman, Boston, 1984.
11. ———, *Inverse Problems in the Mathematical Sciences*, Vieweg, Braunschweig, 1993.
12. ———, *Stable Approximate Evaluation of Unbounded Operators*, Springer, Berlin, 2007.
13. P. C. Hansen, *Regularization tools: a Matlab package for analysis and solution of discrete ill-posed problems*, Numerical Algorithms **6** (1994), 1–35.
14. M. Hegland, *Variable Hilbert scales and their interpolation inequalities with application to Tikhonov regularization*, Appl. Anal. **59** (1995), 207–223.
15. M. Hoffmann and M. Reiss, *Nonlinear estimation for linear inverse problems with error in the operator*, Ann. Stat. **36** (2008), 310–336.
16. B. Hofmann, *Optimization aspects of the generalized discrepancy principle in regularization*, Optimization **17** (1986), 305–316.
17. T. Hohage and M. Pricop, *Nonlinear Tikhonov regularization in Hilbert scales for inverse boundary value problems with random noise*, Inverse Problems and Imaging **2** (2008), 271–290.
18. E. Koptelova, E. Shimanovskaya, A. Artamonov, A. Sazhin, V. Yagola, A. Bruevich, and O. Burkhonov, *Image reconstruction technique and optical monitoring of the QSO2237+0305 from Maidanak Observatory in 2002–2003*, Mon. Not. R. Astron. Soc. **356** (2005), 323–330.
19. S. Krein and Y. I. Petunin, *Scales of Banach spaces*, Russian Math. Surveys **21** (1966), 85–159.
20. A. S. Leonov, *Numerical piecewise-uniform regularization for two-dimensional ill-posed problems*, Inverse Problems **15** (1999), 1165–1176.

21. F. Liu and M. Z. Nashed, *Tikhonov regularization of nonlinear ill-posed poblems with closed operators in Hilbert scales*, J. Inv. Ill-Posed Problems **5** (1997), 363–376.
22. S. Lu, S. V. Pereverzev, and U. Tautenhahn, *Dual regularized total least squares and multi-parameter regularization*, Comput. Meth. Appl. Math. **8** (2008), 253–262.
23. ———, *A model function method in total least squares*, Tech. Report 2008-18, Johann Radon Institute for Computational and Applied Mathematics, 2008.
24. ———, *Regularized total least squares: computational aspects and error bounds*, SIAM J. Matrix Anal. **31** (2009), 918–941.
25. B. A. Mair, *Tikhonov regularization for finitely and infinitely smoothing operators*, SIAM J. Math. Anal. **25** (1994), 135–147.
26. C. Marteau, *Regularization of inverse problems with unknown operator*, Math. Methods Stat. **15** (2006), 415–443.
27. P. Mathé and U. Tautenhahn, *Interpolation in variable Hilbert scales with application to inverse problems*, Inverse Problems **22** (2006), 2271–2297.
28. ———, *Error bounds for regularization methods in Hilbert scales by using operator monotonicity*, Far East J. Math. Sci. **24** (2007), 1–21.
29. F. Mazzone, J. Coyle, A. M. Massone, and M. Piana, *FIST: A fast visualizer for fixed-frequency acoustic and electromagnetic inverse scattering problems*, Simulation Modelling Practice and Theory **14** (2006), 177–187.
30. J. J. Moré and D. C. Sorensen, *Computing a trust region step*, SIAM J. Sci. Stat. Comput. **4** (1983), 553–572.
31. V. A. Morozov, *Regularization Methods for Ill-Posed Problems*, CRC Press, Florida, 1993.
32. M. T. Nair, S. V. Pereverzev, and U. Tautenhahn, *Regularization in Hilbert scales under general smoothing conditions*, Inverse Problems **21** (2005), 1851–1869.
33. F. Natterer, *Error bounds for Tikhonov regularization in Hilbert scales*, Appl. Anal. **18** (1984), 29–37.
34. A. Neubauer, *An a-posteriori parameter choice for Tikhonov-regularization in Hilbert scales leading to optimal convergence rates*, SIAM J. Numer. Anal. **25** (1988), 1313–1326.
35. ———, *Tikhonov regularization of nonlinear ill-posed problems in Hilbert scales*, Appl. Anal. **46** (1992), 59–72.
36. P. P. Mojabi and J. LoVetri, *Adapting the normalized cumulative periodogram parameter-choice method to the Tikhonov regularization of 2-D/TM electromagnetic inverse scattering using Born iterative method* , Progress In Electromagnetics Research M **1** (2008), 111–138.
37. R. A. Renaut and H. Guo, *Efficient algorithms for solution of regularized total least squares*, SIAM J. Matrix Anal. Appl. **26** (2005), 457–476.
38. T. Schröter and U. Tautenhahn, *Error estimates for Tikhonov regularization in Hilbert scales*, Num. Funct. Anal. and Optimiz. **15** (1994), 155–168.
39. D. Sima, S. V. Huffel, and G. H. Golub, *Regularized total least squares based on quadratic eigenvalue problem solvers*, BIT Numerical Mathematics **44** (2004), 793–812.
40. U. Tautenhahn, *Optimal parameter choice for Tikhonov regularization in Hilbert scales*, Inverse Problems in Mathematical Physics, Lecture Notes in Phys. 422 (Berlin) (L. Päivärinta and E. Somersalo, eds.), Springer, 1993, pp. 242–250.
41. ———, *Error estimates for regularization methods in Hilbert scales*, SIAM J. Numer. Anal. **33** (1996), 2120–2130.
42. ———, *On a general regularization scheme for nonlinear ill-posed problems: II. Regularization in Hilbert scales*, Inverse Problems **14** (1998), 1607–1616.
43. ———, *Regularization of linear ill-posed problems with noisy right hand side and noisy operator*, J. Inv. Ill-Posed Problems **16** (2008), 507–523.
44. A. N. Tikhonov and V. Y. Arsenin, *Solution of Ill-Posed Problems*, Wiley, New York, 1977.
45. A. N. Tikhonov, A. S. Leonov, and A. G. Yagola, *Nonlinear Ill-Posed Problems vol. 1 & 2*, Chapman & Hall, London, 1998.
46. G. M. Vainikko, *The discrepancy principle for a class of regularization methods*, USSR Comput. Math. Math. Phys. **22** (1982), 1–19.

47. ______, *The critical level of discrepancy in regularization methods*, USSR Comput. Math. Math. Phys. **23** (1983), 1–9.
48. G. M. Vainikko and A. Y. Veretennikov, *Iteration Procedures in Ill-Posed Problems*, Nauka, Moscow, 1986, In Russian.
49. V. V. Vasin, *Some tendencies in the Tikhonov regularization of ill-posed problems*, J. Inv. Ill-Posed Problems **14** (2006), 813–840.

S. Lu, Johann Radon Institute for Computational and Applied Mathematics, Austrian Academy of Sciences, Altenbergstrasse 69, 4040 Linz, Austria
E-mail address: `shuai.lu@oeaw.ac.at`

S. Pereverzev, Johann Radon Institute for Computational and Applied Mathematics, Austrian Academy of Sciences, Altenbergstrasse 69, 4040 Linz, Austria
E-mail address: `sergei.pereverzyev@oeaw.ac.at`

Y. Shao, Department of Mathematics, Chemnitz University of Technology, 09107 Chemnitz, Germany
E-mail address: `yuanyuan.shao@s2009.tu-chemnitz.de`

U. Tautenhahn, Department of Mathematics, University of Applied Sciences Zittau/Görlitz, P.O.Box 1455, 02755 Zittau, Germany
E-mail address: `u.tautenhahn@hs-zigr.de`