Mathematical Modelling and Analysis
Volume 14 Number 4, 2009, pages 451–466
Doi:10.3846/1392-6292.2009.14.451-466

ISSN 1392-6292 print
ISSN 1648-3510 online
www.vgtu.lt/mma/

A Unified Approach for Regularizing Discretized Linear Ill-Posed Problems

T. Hein

Chemnitz University of Technology
09107 Chemnitz, Germany
E-mail: `torsten.hein@mathematik.tu-chemnitz.de`

Received November 5, 2008; revised October 7, 2009; published online November 10, 2009

Abstract. In this paper we deal with regularization approaches for discretized linear ill-posed problems in Hilbert spaces. As opposite to other contributions concerning this topic the smoothness of the unknown solution is measured with so-called approximative source conditions. This idea allows us to generalize known convergence rates results to arbitrary classes of smoothness conditions including logarithmic and general source conditions. The considerations include an a-posteriori parameter choice strategy for the regularization parameter and the discretization level. Results of one numerical example are presented.

Key words: inverse problem, linear ill-posed problem,regularization, projection method, convergence rates.

1 Introduction

Let $\mathcal{X}$ and $\mathcal{Y}$ denote Hilbert spaces and let $A : \mathcal{X} \longrightarrow \mathcal{Y}$ describe a linear and compact operator with non-closed range, i.e. $\mathcal{R}(A) \neq \overline{\mathcal{R}(A)}$. We introduce the notation $a := \|A\|$. We consider convergence rates for solving the linear ill-posed operator equation

$$A\,x = y^\delta, \quad x \in \mathcal{X}, \;\; y^\delta \in \mathcal{Y}, \tag{1.1}$$

approximately for given noisy data $y^\delta \in \mathcal{Y}$. Let $y = y^0$ be the exact data and $\delta \geq 0$ denotes the noise level, i.e. $\|y - y^\delta\| \leq \delta$. Moreover, we assume the existence of $x^\dagger := A^\dagger y$, which is referred to as the exact solution of (1.1) for given exact data y. Here $A^\dagger$ is the Moore-Penrose inverse of A. In order to solve (1.1) numerically we replace equation (1.1) by

$$Q_h A P_h\, x = Q_h y^\delta, \quad x \in \mathcal{X}, \quad y^\delta \in \mathcal{Y}, \tag{1.2}$$

where $P_h : \mathcal{X} \longrightarrow \mathcal{X}$ and $Q_h : \mathcal{Y} \longrightarrow \mathcal{Y}$ are orthogonal projections satisfying

$$\|A(I - P_h)\| \leq \xi_h \qquad \text{and} \qquad \|(I - Q_h)A\| \leq \eta_h$$

for given bounds ξ_h and η_h. This kind of definition for the bounds of the discretization errors has been well-established in inverse problems, see e.g. [18] and [20]. We set $A_h := Q_h A P_h$. As usual, h describes the discretization level and we assume $\xi_h \to 0$ and $\eta_h \to 0$ for $h \to 0$.

Of course, if $\mathcal{R}(P_h)$ and $\mathcal{R}(Q_h)$ are finite-dimensional, equation (1.2) is not ill-posed anymore. On the other hand, (1.2) becomes more and more ill-conditioned the smaller we choose h. Finite dimensional approximation in combination with a choice of the discretization level h depending on the noise level δ is well-known as regularization by projection, see [18] and [2, Chapter 7]. We also refer to [12] to some new results and further references therein. However, numerous numerical studies indicate that we can often apply only a coarse discretization which is insufficient for problems arising in practical applications. Hence, it makes sense to introduce an additional regularization strategy $\{g_\alpha\}$, $0 < \alpha \leq a^2$, i.e. we replace (1.2) by calculating

$$x^\delta_{\alpha,h} := g_\alpha(A_h^* A_h) A_h^* Q_h y^\delta = g_\alpha(A_h^* A_h) A_h^* y^\delta. \tag{1.3}$$

In order to derive convergence rates $x^\delta_{\alpha,h} \to x^\dagger$ for $\delta \to 0$ (respectively $\alpha = \alpha(\delta) \to 0$ and $h = h(\delta) \to 0$) we need additional (smoothness) conditions to the exact solution $x^\dagger$. They are usually given by a general source condition

$$x^\dagger = \varphi(A^*A)\omega, \qquad \omega \in \mathcal{X}, \tag{1.4}$$

for an index function $\varphi(t)$, $t \geq 0$. In this context, a continuous and strictly increasing function $\varphi(t)$, $t \geq 0$, is called an index function if $\varphi(0) = 0$. The original idea of power-type functions $\varphi(t) = t^\nu$, $t \geq 0$, for some $\nu > 0$ is well-studied in the undiscretized situation, see e.g. [4] and the references therein. Later on, in [17] and [22] the idea was generalized to more general index functions $\varphi(t)$, $t \geq 0$, see also [16] for some further results. Moreover, for each $x^\dagger \in \mathcal{R}(A^\dagger)$ we can find an index function $\varphi(t)$, $t \geq 0$, such that (1.4) holds, see [10].

If we additionally consider the discretization aspect further difficulties occur in deriving convergence rates. In particular, we have to find an estimate for the term $\|\varphi(A_h^* A_h) - \varphi(A^*A)\|$ which is not trivial job for arbitrary index functions, see [15]. However, such an estimate can be found for power-type index functions. This situation was treated in [20] in detail. In [15] more general index functions were considered. But unfortunately, the achieved results therein are still restricted to some limited classes of index functions.

Therefore we follow an alternative strategy. We refer to (1.4) as reference or benchmark source condition which we allow to be violated. Then we replace the source condition (1.4) by a so-called approximate source condition

$$x^\dagger = \varphi(A^*A)\omega + \upsilon, \qquad \omega, \upsilon \in \mathcal{X}, \tag{1.5}$$

introduced in [7] and [8]. Here we assume, that the function $\varphi(t)$, $t \geq 0$, is chosen such that $x^\dagger \notin \mathcal{R}(\varphi(A^*A))$. On the other hand, we have

$$x^\dagger \in \mathcal{R}(A^\dagger) \subset \overline{\mathcal{R}(A^\dagger)} = \mathcal{N}(A)^\perp = \overline{\mathcal{R}((A^*A)^{\frac{1}{2}})} = \overline{\mathcal{R}(\varphi(A^*A))}$$

by definition of the Moore-Penrose inverse $A^\dagger$. However, this indicates, that for arbitrary $d > 0$ we can find elements $\omega, \upsilon \in \mathcal{X}$ with $\|\upsilon\| \leq d$ satisfying (1.5). Introducing distance functions, this observation allows us to present convergence rates in the undiscretized situation. This is a topic of Section 2. Then, the idea is quite simple: we choose a power-type index $\varphi(t) = t^\nu$, $t \geq 0$, with sufficiently large exponent $\nu > 0$ and apply the results of [20] for finding estimates for the critical terms. The additional terms which are caused by the violation of the power-type source conditions can be treated separately. We will show this in Section 3. So we can extend and improve the known convergence rates results of [15, 20] to arbitrary classes of smoothness conditions on the exact solution $x^\dagger$. In particular, these considerations also includes logarithmic source conditions introduced in [11].

Moreover, the presented analysis does not depend on any type of source condition. So discretization and regularization strategy under consideration can be applied to arbitrary $x^\dagger \in \mathcal{R}(A^\dagger)$. This is the main purpose of the present paper.

Therefore, the paper is organized as follows: in Section 2 approximative source conditions are opposed to general source conditions in the undiscretized situation. In both cases results on error bounds for the regularized solutions of (1.1) were summarized and a first statement about the correlation of both approaches was presented. In Section 3 we consider error bounds for the term $\|x^\delta_{\alpha,h} - x^\dagger\|$ depending on δ, α and the approximative source condition. There, a choice of the discretization levels h depending on the regularization parameter α is suggested. Section 4 is devoted to a convergence rates result based on an a-priori parameter choice strategy $\alpha = \alpha(\delta)$ whereas following an a-posteriori choice according the balancing principle of Lepskij is proposed. Finally we illustrate these theoretical results by a numerical example.

2 General vs. Approximate Source Conditions

In order to verify convergence rates for linear (and nonlinear) regularization approaches two concepts have been established in the recent years. We briefly recall both ideas. Introducing further notations we recall the definition of a regularization $\{g_\alpha\}$, $0 < \alpha \leq a^2$, see [16] and [9].

Definition 1. A family $\{g_\alpha\}$, $0 < \alpha \leq a^2 = \|A^*A\|$, of piecewise continuous functions is called regularization if there exist constants $C_1 > 0$ and $C_2 > 0$ such that for $0 < \alpha \leq a^2$

$$\sup_{0<t\leq a^2} \sqrt{t}\,|g_\alpha(t)| \leq \frac{C_1}{\sqrt{\alpha}} \quad \text{and} \quad \sup_{0<t\leq a^2} |1 - t\,g_\alpha(t)| \leq C_2. \tag{2.1}$$

For obtaining estimates of the regularization error we need the concept of (general) qualifications, see also [16, Definition 1].

Definition 2. The regularization $\{g_\alpha\}$ is said to have qualification $\varphi(t)$, $t \geq 0$, for an index function $\varphi(t)$, if there exists a constant $C_\varphi > 0$ such that

$$\sup_{0<t\leq a^2} |1 - t\,g_\alpha(t)|\varphi(t) \leq C_\varphi \varphi(\alpha), \quad 0 < \alpha \leq a^2.$$

First we deal with general source conditions. For given index function $\varphi(t)$, $t \geq 0$, we assume that $x^\dagger$ satisfies the source condition (1.4) with $\|\omega\| \leq R$ for some $R > 0$. Then, if this function $\varphi(t)$, $t \geq 0$ is a qualification of the regularization $\{g_\alpha\}$, we obtain the error estimate

$$\|x_\alpha^\delta - x^\dagger\| \leq \frac{C_1}{\sqrt{\alpha}}\delta + C_\varphi \varphi(\alpha) R, \qquad 0 < \alpha \leq a^2, \tag{2.2}$$

where $x_\alpha^\delta := g_\alpha(A^*A)A^*y^\delta$ denotes the regularized solution of equation (1.1), see [16]. In the following we always suppose that the index function $\varphi(t)$, $t \geq 0$ is a qualification of the regularization $\{g_\alpha\}$.

Alternatively we can assume, that a source condition (1.4) with given index function $\varphi(t)$, $t \geq 0$, is violated. Therefore we introduce the sets

$$\mathcal{M}_\varphi(R,d) := \{x \in \mathcal{X} \,:\, x = \varphi(A^*A)\omega + \upsilon, \ \|\omega\| \leq R, \ \|\upsilon\| \leq d\}, \quad R,\, d \geq 0.$$

Then, for $x^\dagger \in \mathcal{M}_\varphi(R,d)$, i.e. $x^\dagger = \varphi(A^*A)\omega + \upsilon$ with $\omega, \upsilon \in \mathcal{X}$ satisfying $\|\omega\| \leq R$ and $\upsilon \leq d$, we conclude from the classical linear regularization theory

$$\begin{aligned}
\|x_\alpha^\delta - x^\dagger\| &= \|g_\alpha(A^*A)A^*(y^\delta - y + y) - x^\dagger\| \\
&\leq \|(I - g_\alpha(A^*A)A^*A)x^\dagger\| + \|g_\alpha(A^*A)A^*(y^\delta - y)\| \\
&= \|(I - g_\alpha(A^*A)A^*A)\left[\varphi(A^*A)\omega + \upsilon\right]\| + \|g_\alpha(A^*A)A^*(y - y^\delta)\| \\
&\leq C_\varphi \varphi(\alpha) R + C_2 d + \frac{C_1}{\sqrt{\alpha}}\delta.
\end{aligned}$$

In order to prove convergence rates based on approximate source conditions we additionally need the concept of distance functions. The idea of approximate source conditions was originally introduced in [2, Theorem 6.8] for measuring the approximation term $\|x - A^*\tilde{\omega}\|$ for given $x \notin \mathcal{R}(A^*)$ with $\tilde{\omega} \in \mathcal{Y}$, $\|\tilde{\omega}\| \leq R$ for each $R > 0$. Generalizing the idea, we recall the definition of distance functions, see e.g. [3].

Definition 3. For given $x \in \mathcal{X}$ and index function $\varphi(t)$, $t \geq 0$, the distance function $d_\varphi(\cdot; x) : [0, \infty) \longrightarrow \mathbb{R}$ is defined as

$$d_\varphi(R; x) := \min\left\{\|x - \varphi(A^*A)\omega\| \,:\, \omega \in \mathcal{X}, \ \|\omega\| \leq R\right\}, \quad R \geq 0.$$

Note, that the nonnegative function $d_\varphi(R;x)$ is well-defined for each $x \in \mathcal{X}$, i.e., for each $R \geq 0$ there exists an element $\omega = \omega(R)$ with $d_\varphi(R;x) = \|x - \varphi(A^*A)\omega\|$ and $\|\omega\| \leq R$, see [23, Theorem 38.A]. The distance functions are non-increasing with $d_\varphi(R;x) \to 0$ for $R \to \infty$ if $x \in \overline{\mathcal{R}(\varphi(A^*A))} = \mathcal{R}(A^\dagger)$ since

$$\overline{\mathcal{R}(\varphi(A^*A))} = \overline{\mathcal{R}((A^*A)^{\frac{1}{2}})} = \overline{\mathcal{R}(A^*)} = \mathcal{N}(A)^\perp = \mathcal{R}(A^\dagger).$$

We have $d_\varphi(R;x) > 0$ for all $R \geq 0$ if $x \notin \mathcal{R}(\varphi(A^*A))$ and $d_\varphi(R;x) = 0$ for all $R \geq \|\omega\|$ if $x = \varphi(A^*A)\omega$ and $\|\omega\| = R$. Altogether, the distance function $d_\varphi(R;x)$ gives us a quantity for measuring the violation of the general source condition $x \in \mathcal{R}(\varphi(A^*A))$, see also [7].

With the aid of the distance function, the error bound for approximative source conditions has the same structure as in the case of general source conditions. A similar result is presented in [3, Theorem 2.5] for Tikhonov regularization with a different notation, see also [9, Theorem 5.5] for the case $\delta = 0$.

Theorem 1. *Let the index function $\varphi(t)$, $t \geq 0$, be a qualification of the regularization $\{g_\alpha\}$. We assume $x^\dagger \notin \mathcal{R}(\varphi(A^*A))$ has distance function $d(R) := d_\varphi(R; x^\dagger)$. Let $\Theta(R) := \varphi^{-1}(d(R)R^{-1})$. Then the estimate*

$$\|x_\alpha^\delta - x^\dagger\| \leq \frac{C_1}{\sqrt{\alpha}}\delta + (C_2 + C_\varphi)\, d\left(\Theta^{-1}(\alpha)\right) \tag{2.3}$$

holds. Moreover, we define the functions

$$\Psi(\alpha) := \sqrt{\alpha} d(\Theta^{-1}(\alpha)), \quad \Phi(R) := \sqrt{\Theta^{-1}(d(R)R^{-1})} d(R).$$

Then, an a-priori parameter choice $\alpha := \Psi^{-1}(\delta)$ leads to a convergence rate

$$\|x_\alpha^\delta - x^\dagger\| \leq (C_1 + C_2 + C_\varphi)\, d\left(\Phi^{-1}(\delta)\right). \tag{2.4}$$

Proof. We observe that $x^\dagger \in \overline{\mathcal{R}(\varphi(A^*A))}$. By definition, $x^\dagger \in \mathcal{M}_\varphi(R, d(R))$ for all $R \geq 0$. In particular, $x^\dagger = \varphi(A^*A)\omega_R + \upsilon_R$ with $\|\omega_R\| \leq R$ and $\|\upsilon_R\| = d(R)$ holds for all $R > 0$. Hence, the estimate

$$\|x_\alpha^\delta - x^\dagger\| \leq \frac{C_1}{\sqrt{\alpha}}\delta + C_\varphi \varphi(\alpha) R + C_2 d(R)$$

holds for all $R \geq 0$. We choose $R = R(\alpha) > 0$ such that

$$R\,\varphi(\alpha) = d(R) \;\Leftrightarrow\; \varphi(\alpha) = \frac{d(R)}{R} \;\Leftrightarrow\; \alpha = \varphi^{-1}\left(d(R)R^{-1}\right) = \Theta(R).$$

This proves (2.3). The choice $\alpha = \Psi^{-1}(\delta)$ implies

$$\frac{\delta}{\sqrt{\alpha}} = d\left(\Theta^{-1}(\alpha)\right) \;\Leftrightarrow\; \delta = \sqrt{\alpha} d\left(\Theta^{-1}(\alpha)\right).$$

On the other hand, we have

$$\frac{\delta}{\sqrt{\alpha}} = d(R) \;\Leftrightarrow\; \delta = d(R)\sqrt{\alpha} = \sqrt{\Theta^{-1}(d(R)R^{-1})} d(R) = \Phi(R).$$

The choice $R = \Phi^{-1}(\delta)$ leads to (2.4). □

Since $d(R)$ is non-increasing the function $R \mapsto d(R)/R$ is strictly decreasing on $(0, \infty)$ with $d(R)/R \to \infty$ for $R \to 0$ and $d(R)/R \to 0$ for $R \to \infty$. This implies, that $\Theta(R)$ is strictly decreasing. Hence, $\Theta^{-1}(\alpha)$ is well-defined and strictly decreasing. Moreover, we observe that $d(\Theta^{-1}(\alpha))$ is an index function, i.e. it is increasing with $d(\Theta^{-1}(\alpha)) \to 0$ for $\alpha \to 0$. Analogous calculations show that the functions $\Psi^{-1}(\delta)$ and $\Phi^{-1}(\delta)$ are well-defined.

Comparing both error bounds (2.2) and (2.3) we notice similar structure. In both cases the term $C_1\,\delta/\sqrt{\alpha}$ is based on the error in the given data. The second term is the regularization error which depends on the regularization parameter α and the (approximative) source condition, which is usually unknown. We use this structure later for proposing an a-posteriori parameter choice rule based on the balancing principle [13].

Studying the connection between general and approximate source conditions is topic of recent research. In [3, Theorem 3.1 and Corollary 3.3] we find the following result for power-type reference source conditions. We also refer to [9, Theorem 5.9] for a more general result.

Proposition 1. *Let $\mathcal{X}$ and $\mathcal{Y}$ be Hilbert spaces and $A \in L(\mathcal{X},\mathcal{Y})$ be injective and compact. Suppose $\varphi(t) = t^\mu$, $t \geq 0$, for some exponent $\mu > 0$. Assume $x^\dagger \notin \mathcal{R}(\varphi(A^*A))$ has distance function $d_\varphi(R;x^\dagger)$, $R \geq 0$.*

(i) If $x^\dagger \in \mathcal{R}\left((A^*A)^\nu\right)$ for some exponent $0 < \nu < \mu$ then the estimate

$$d_\varphi(R;x^\dagger) \leq K\,R^{\frac{\nu}{\nu-\mu}}, \quad R > 0, \tag{2.5}$$

holds for some constant $K > 0$.

(ii) If the distance function $d_\varphi(R,x^\dagger)$ satisfies (2.5) for all $R > 0$ then $x^\dagger \in \mathcal{R}\left((A^*A)^{\tilde{\nu}}\right)$ holds for all $0 < \tilde{\nu} < \nu$.

Remark 1. Let $\varphi(t) = \sqrt{t}$, $t \geq 0$, be a qualification of $\{g_\alpha\}$. Then, for $x^\dagger \in \mathcal{R}((A^*A)^\nu)$ for some $0 < \nu < \frac{1}{2}$ and $x_\alpha = R_\alpha y$ we have a regularization error $\|x_\alpha - x^\dagger\| \leq C\,\alpha^\nu$. On the other hand, the supposed source condition implies that $d(R) = K\,R^{\frac{2\nu}{2\nu-1}}$ is an upper bound for the distance function. We therefore assume $d_\varphi(R;x^\dagger) = K\,R^{\frac{2\nu}{2\nu-1}}$. Then from Theorem 1 we deduce with some generic constant $\mathcal{C} > 0$ that $\Theta(R) = \mathcal{C}\,R^{\frac{2}{2\nu-1}}$ and $d_\varphi(\Theta^{-1}(\alpha);x^\dagger) = \mathcal{C}\,\alpha^\nu$. Hence, the error estimates (2.2) and (2.3) are of the same order. This example shows close connection of both approaches of (general) source condition and approximative source condition in this specific situation.

We also remark that we cannot conclude from (2.5) that $x^\dagger \in \mathcal{R}\left((A^*A)^\nu\right)$. This is a consequence of the well-known converse result [19, Corollary 2.6]. Moreover, this gap coincides with an observation in [6, Corollary 1]: we cannot formulate a maximal source condition (1.4) for the element $x^\dagger \in \mathcal{X}$. We present an example which shows that this gap really occurs.

Example 1. We set $\mathcal{X} = \mathcal{Y} = L^2(0,1)$ and consider the linear operator $A : L^2(0,1) \longrightarrow L^2(0,1)$ given by

$$(A\,x)(t) := \int_0^t x(\tau)\,d\tau, \qquad t \in [0,1],\ x \in L^2(0,1).$$

Moreover we assume $x^\dagger \equiv 1$. From [8, Example 4] we know that $x^\dagger \in \mathcal{R}\left((A^*A)^\nu\right)$ for all $\nu < \frac{1}{4}$ but $x^\dagger \notin \mathcal{R}\left((A^*A)^{\frac{1}{4}}\right)$. We set $\varphi(t) := \sqrt{t}$, $t \geq 0$.

Since $\mathcal{R}\left((A^*A)^{\frac{1}{2}}\right) = \mathcal{R}(A^*)$ we can find in [6, Example 4.5] an estimate for the distance function $d_\varphi(R; x^\dagger)$. It is given by $d_\varphi(R; x^\dagger) \le \tilde{K}\, R^{-1} = \tilde{K}\, R^{\frac{\tilde{\nu}}{\tilde{\nu}-1/2}}$, $R \ge 1$ with $\tilde{\nu} = \frac{1}{4}$. Moreover, the construction of the distance function shows that we can generalize this result to arbitrary function x satisfying $x' \in L^2(0,1)$ and $x(1) \neq 0$. Comparing both results shows the following: the estimate of the distance function is of correct order, i.e. the exponent cannot be decreased. Moreover, based on distance functions we can prove for $x^\dagger$ the optimal convergence rates $\|x_\alpha^\delta - x^\dagger\| = \mathcal{O}\left(\delta^{\frac{1}{3}}\right)$ as $\delta \to 0$ by choosing $\alpha := \delta^{\frac{4}{3}}$. This rate cannot be proved with the concept of general source conditions.

3 Error Bounds Based on Approximate Source Conditions

According to (1.3) we introduce the operator $R_{\alpha,h} := g_\alpha(A_h^* A_h) A_h^*$ and set

$$x_{\alpha,h}^\delta := R_{\alpha,h} Q_h y^\delta = g_\alpha(A_h^* A_h) A_h^* Q_h y^\delta = g_\alpha(A_h^* A_h) A_h^* y^\delta.$$

Then we derive

$$\begin{aligned} x^\dagger - R_{\alpha,h} y^\delta &= (I - g_\alpha(A_h^* A_h) A_h^* A_h)\, x^\dagger + g_\alpha(A_h^* A_h) A_h^* (A_h x^\dagger - Q_h y^\delta) \\ &= (I - g_\alpha(A_h^* A_h) A_h^* A_h)\, x^\dagger + g_\alpha(A_h^* A_h) A_h^* Q_h (y - y^\delta) \\ &\quad + g_\alpha(A_h^* A_h) A_h^* Q_h A (P_h - I) x^\dagger. \end{aligned}$$

The second term we estimate by

$$\left\| g_\alpha(A_h^* A_h) A_h^* Q_h (y - y^\delta) \right\| \le C_1 \frac{\delta}{\sqrt{\alpha}}.$$

An error bound for $\|x_{\alpha,h}^\delta - x^\dagger\|$ now depends on the specific (approximative) source condition. Therefore we assume that the regularization $\{g_\alpha\}$ has qualification $\varrho(t)$, $t \ge 0$, which covers the index function $\varphi(t)$, $t \ge 0$. We recall, that the function $\varrho(t)$, $t \ge 0$, covers $\varphi(t)$, $t \ge 0$, if $\varrho(\alpha)/\varphi(\alpha)$ remains bounded as $\alpha \to 0$. Moreover, let $x^\dagger \in \mathcal{M}_\varphi(R, d)$ hold for some $R \ge 0$ and $d \ge 0$. Hence, there exist $\omega, \upsilon \in \mathcal{X}$ with $\|\omega\| \le R$ and $\|\upsilon\| \le d$ such that $x^\dagger = \varphi(A^* A)\omega + \upsilon$. Starting with the first part we conclude

$$\begin{aligned} \left\| (I - g_\alpha(A_h^* A_h) A_h^* A_h)\, x^\dagger \right\| &= \| (I - g_\alpha(A_h^* A_h) A_h^* A_h) \left[\varphi(A^* A)\omega + \upsilon \right] \| \\ &\le \| (I - g_\alpha(A_h^* A_h) A_h^* A_h)\, \varphi(A^* A) \|\; R + C_2\, d. \end{aligned}$$

Moreover, for the term $E := \|(I - g_\alpha(A_h^* A_h) A_h^* A_h)\, \varphi(A^* A)\|$ we can find the estimate

$$\begin{aligned} E &\le \| \big(I - g_\alpha(A_h^* A_h) A_h^* A_h\big) \varphi(A_h^* A_h) \| + \left\| \big(I - g_\alpha(A_h^* A_h) A_h^* A_h\big) \big(\varphi(A^* A) \right. \\ &\quad \left. - \varphi(A_h^* A_h)\big) \right\| = \left\| \big(I - g_\alpha(A_h^* A_h) A_h^* A_h\big) \varphi(A_h^* A_h) \right\| + \left\| \varphi(A^* A) - P_h \varphi(A^* A) \right. \\ &\quad \left. + P_h \big(\varphi(A^* A) - \varphi(A_h^* A_h)\big) - g_\alpha(A_h^* A_h) A_h^* A_h P_h \big(\varphi(A^* A) - \varphi(A_h^* A_h)\big) \right\| \\ &\le \| (I - g_\alpha(A_h^* A_h) A_h^* A_h)\, \varphi(A_h^* A_h) \| + \| (I - P_h) \varphi(A^* A) \| \\ &\quad + \| (I - g_\alpha(A_h^* A_h) A_h^* A_h)\, (P_h \varphi(A^* A) - \varphi(A_h^* A_h)) \| \\ &\le C_\varphi \varphi(\alpha) + C_2\, \| P_h \varphi(A^* A) - \varphi(A_h^* A_h) \| + \| (I - P_h) \varphi(A^* A) \| \end{aligned}$$

by using $P_h\varphi(A_h^*A_h) = \varphi(A_h^*A_h)$. The latter is evident since $\mathcal{N}(P_h) \subseteq \mathcal{N}(A_h)$ and

$$\mathcal{R}(\varphi(A_h^*A_h)) = \mathcal{R}((A_h^*A_h)^{\frac{1}{2}}) = \mathcal{R}(A_h^*) = \mathcal{N}(A_h)^\perp \subseteq \mathcal{N}(P_h)^\perp .$$

Finally, we can derive the estimate

$$\begin{aligned}\big\|g_\alpha(A_h^*A_h)A_h^*Q_hA(P_h-I)x^\dagger\big\| &\le \frac{C_1}{\sqrt{\alpha}}\big(\big\|Q_hA(P_h-I)\varphi(A^*A)\big\|R+\xi_h d\big)\\ &= \frac{C_1}{\sqrt{\alpha}}\big(\big\|Q_hA(P_h-I)(P_h-I)\varphi(A^*A)\big\|R+\xi_h d\big)\\ &\le \frac{C_1\xi_h}{\sqrt{\alpha}}\left(\|(P_h-I)\varphi(A^*A)\|\, R+d\right).\end{aligned}$$

It is difficult to find optimal estimates for the terms $\|(I-P_h)\varphi(A^*A)\|$ and $\|P_h\varphi(A^*A) - \varphi(A_h^*A_h)\|$ for an arbitrary function $\varphi(t)$, $t \ge 0$, see e.g. [15] and the discussion therein. On the other hand, for $\varphi(t) := t^\nu$, $\nu > 0$, the following result has been established, see [20, Lemma 4.3. and 4.4.].

Proposition 2. *Suppose that $A \in L(\mathcal{X},\mathcal{Y})$. For $\nu > 0$ and orthogonal projections $P \in L(\mathcal{X})$ and $Q \in L(\mathcal{Y})$ we obtain*

$$\|(I-P)(A^*A)^\nu\| \le \hat{C}_\nu \|A(I-P)\|^{\min\{2\nu,1\}}$$

where $\hat{C}_\nu = 1$ for $\nu \le \frac{1}{2}$ and $\hat{C}_\nu = \|A\|^{2\nu-1}$ for $\nu > \frac{1}{2}$ as well as

$$\|P(A^*A)^\nu - (Q\,A\,P)^\nu\| \le \tilde{C}_\nu\left(\|A(I-P)\|^{\min\{2\nu,1\}} + \|(I-Q)A\|^{\min\{2\nu,2\}}\right).$$

The mapping $\nu \mapsto \tilde{C}_\nu$ is bounded in $(0,\nu_0]$ for any $\nu_0 > 0$.

By applying Proposition 2 we obtain the following result.

Theorem 2. *Let $\{g_\alpha\}$ be a regularization with qualification $\varphi(t) = t^\nu$, $t \ge 0$, for some $\nu \ge \frac{1}{2}$. Suppose $x^\dagger \in \mathcal{M}_\varphi(R,d)$ for some $R,d \ge 0$. Then*

$$\begin{aligned}\|x_{\alpha,h}^\delta - x^\dagger\| \le{}& C_1\frac{\delta}{\sqrt{\alpha}} + R\left(C_\varphi\alpha^\nu + C_2(\hat{C}_\nu+\tilde{C}_\nu)\xi_h + \tilde{C}_\nu\eta_h^{\min\{2\nu,2\}}\right)\\ &+\frac{C_1}{\sqrt{\alpha}}\left(\xi_h d + \hat{C}_\nu\xi_h^2 R\right) + C_2 d.\end{aligned} \tag{3.1}$$

Estimate (3.1) now provides a balanced choice of the discretization level h with respect to the bounds ξ_h and η_h depending on the regularization parameter α. We present the following first main result.

Theorem 3. *Let $\{g_\alpha\}$ be a regularization with qualification $\varphi(t) = t^\nu$, $t \ge 0$, for some $\nu \ge \frac{1}{2}$. Suppose $x^\dagger \in \mathcal{M}_\varphi(R,d)$ for some $R,d \ge 0$. If $\alpha \le 1$, then for the choice $\xi_h \le C\alpha^\nu$ and $\eta_h \le C\alpha^{\frac{\nu}{\min\{2\nu,2\}}}$ we obtain the error bound*

$$\|x_{\alpha,h}^\delta - x^\dagger\| \le C_1\frac{\delta}{\sqrt{\alpha}} + C_3R\,\alpha^\nu + C_4 d \tag{3.2}$$

for three constants $C_1, C_3, C_4 > 0$.

Proof. Since $\nu \geq \frac{1}{2}$ and $\alpha \leq 1$ we have $\frac{\xi_h}{\sqrt{\alpha}} \leq C\alpha^{\nu-\frac{1}{2}} \leq C$ and $\frac{\xi_h^2}{\sqrt{\alpha}} \leq C^2\alpha^{2\nu-\frac{1}{2}} \leq C^2\alpha^{\nu}$. Then, by setting

$$C_3 := C_\varphi + C_2 C(\hat{C}_\nu + \tilde{C}_\nu) + \tilde{C}_\nu C^{\min\{2\nu,2\}} + C_1 C^2 \hat{C}_\nu,\ C_4 := C_2 + C_1 C,$$

the estimate (3.2) follows immediately from the bound (3.1). □

The same idea as in Theorem 1 now leads to the following consequence.

Corollary 1. Let $\{g_\alpha\}$ be a regularization with qualification $\varphi(t) = t^\nu$, $t \geq 0$, for some $\nu \geq \frac{1}{2}$. Assume $x^\dagger \notin \mathcal{R}(\varphi(A^*A))$ has distance function $d(R) := d_\varphi(R; x^\dagger)$. We define $\Theta(R) := (d(R)R^{-1})^{\frac{1}{\nu}}$. If $0 < \alpha \leq \min\{1, a^2\}$, then for the choice $\xi_h \leq C\alpha^\nu$ and $\eta_h \leq C\alpha^{\frac{\nu}{\min\{2\nu,2\}}}$ we obtain the error bound

$$\|x_{\alpha,h}^\delta - x^\dagger\| \leq C_1 \frac{\delta}{\sqrt{\alpha}} + (C_3 + C_4) d\left(\Theta^{-1}(\alpha)\right). \tag{3.3}$$

We will also point out the following: all additional errors based on the discretization are concentrated in the second term of the bound (3.3). This is quite remarkable and opposite to other approaches, see e.g. [15]. Of course this might lead to a finer discretization. But as we will see, it has advantages by applying the balancing principle [13] for choosing the regularization parameter α. There, only the first (noise) error term plays a role for the choice of α, which only depends on δ, α and C_1. The latter constant is given by the chosen regularization method. In particular, the term does not depend on the specific general or approximative source condition. For the sake of completeness we present the following result, see also [15, Section 4].

Corollary 2. Assume $x^\dagger \in \mathcal{R}((A^*A)^\nu)$ for some $\nu \geq \frac{1}{2}$ and t^ν, $t \geq 0$, is qualification of the regularization $\{g_\alpha\}$. Then, the choice

$$\xi_h \leq C \min\left\{\alpha^\nu, \frac{\delta}{\sqrt{\alpha}}\right\} \quad \text{and} \quad \eta_h^{\min\{2\nu,2\}} \leq C\frac{\delta}{\sqrt{\alpha}}$$

yields an error bound

$$\|x_{\alpha,h}^\delta - x^\dagger\| \leq C_5 \frac{\delta}{\sqrt{\alpha}} + C_6 \alpha^\nu$$

for two positive constants C_5 and C_6.

Note, that in particular the constant C_5 now depends on $R := \|\omega\|$, where $\omega \in \mathcal{X}$ is the element satisfying the source condition (1.4). Hence, for applying the balancing principle, an a-priori information about R is needed.

4 On Convergence Rates

We now present main results concerning convergence rates. First, we suppose that $x^\dagger$ satisfies a source condition (1.4) with a qualification φ of power type.

Proposition 3. *Let $\{g_\alpha\}$ be a regularization with qualification $\varphi(t) = t^\nu$, $t \geq 0$, for some $\nu \geq \frac{1}{2}$ and $\Psi(\alpha) := \varphi(\alpha)\sqrt{\alpha} = \alpha^{\frac{2\nu+1}{2}}$. Moreover, $x^\dagger \in \mathcal{R}((A^*A)^\nu)$. Then the choice $\alpha = \Psi^{-1}(\delta) = \delta^{\frac{2}{2\nu+1}}$, $\xi_h \leq C\alpha^\nu$ and $\eta_h \leq C\alpha^{\frac{\nu}{\min\{2\nu,2\}}}$ for some constant $C > 0$ yields a convergence rate*

$$\|x^\delta_{\alpha,h} - x^\dagger\| = \mathcal{O}\left(\delta^{\frac{2\nu}{2\nu+1}}\right) \quad as \quad \delta \to 0.$$

Proof. By definition, we have $x^\dagger \in \mathcal{M}_\varphi(R, 0)$ for $R > 0$ chosen sufficiently large. Hence, we can apply estimate (3.2) $d = 0$. Choosing $\alpha = \alpha(\delta)$ such that

$$\frac{\delta}{\sqrt{\alpha}} = \alpha^\nu \;\Leftrightarrow\; \alpha = \delta^{\frac{2}{2\nu+1}} = \Psi(\delta) \;\Leftrightarrow\; \|x^\delta_{\alpha,h} - x^\dagger\| \leq (C_1 + C_3 R)\,\delta^{\frac{2\nu}{2\nu+1}},$$

we derive the desired convergence rate result. □

Note, that we get the same rate of convergence as in the case without discretization which is known as the optimal one. Hence, the above suggested discretization strategy does not spoil the convergence rate. We can now present a similar convergence rate result in the case of violated source condition.

Theorem 4. *Let $\{g_\alpha\}$ be a regularization with qualification $\varphi(t) = t^\nu$, $t \geq 0$, for some $\nu \geq \frac{1}{2}$. Moreover, $x^\dagger \notin \mathcal{R}(\varphi(A^*A))$ has distance function $d(R) := d_\varphi(R; x^\dagger)$. We define the functions $\Theta(R) := (d(R)R^{-1})^{1/\nu}$, $\tilde{\Psi}(\alpha) := d(\Theta^{-1}(\alpha))\sqrt{\alpha}$ and $\Phi(R) := \sqrt{\Theta^{-1}(d(R)R^{-1})}d(R)$. Then the choice $\alpha = \tilde{\Psi}^{-1}(\delta)$, $\xi_h \leq C\alpha^\nu$ and $\eta_h \leq C\alpha^{\frac{\nu}{\min\{2\nu,2\}}}$ for some constant $C > 0$ yields a convergence rate*

$$\|x^\delta_{\alpha,h} - x^\dagger\| = \mathcal{O}\left(d\left(\Phi^{-1}(\delta)\right)\right) \quad as \quad \delta \to 0.$$

The proof is essentially the same as the second balancing step in the proof of Theorem 1.

The above considerations leave an open question. Assume the regularization $\{g_\alpha\}$ has no maximal qualification in the classical sense, i.e. $\varphi(t) = t^\nu$, $t \geq 0$, is qualification for each $\nu > 0$. If $x^\dagger \in \mathcal{R}\left((A^*A)^\nu\right)$ for each $\nu > 0$, then we cannot present an optimal convergence rate for the discretized problem. We only state the following sub-optimal convergence rate result.

Corollary 3. Let $\{g_\alpha\}$ be a regularization such that $\varphi(t) = t^\nu$, $t \geq 0$, is a qualification for all $\nu > 0$. Let $x^\dagger \in \mathcal{R}(\varphi(A^*A))$ for an index function φ with $\varphi(t)t^{-\nu} \to 0$ for $t \to 0$ and each $\nu > 0$. Then the choice $\alpha := \tilde{\Psi}^{-1}(\alpha)$ with function $\tilde{\Psi}(\alpha)$ as in Theorem 4, $\xi_h \leq C\varphi(\alpha)$ and $\eta_h \leq C\sqrt{\varphi(\alpha)}$ for some constant $C > 0$ yields a convergence rate

$$\|x^\delta_{\alpha,h} - x^\dagger\| \sim \mathcal{O}\left(\delta^{\frac{2\nu}{2\nu+1}}\right) \quad \text{for} \quad \delta \to 0$$

for each $\nu > 0$.

The proof is trivial by noticing, that the estimate (3.3) holds asymptotically for each $\nu > 0$ with $R = R(\nu) > 0$ and $d = 0$.

5 The Lepskij–Principle

We now present an a-posteriori parameter choice strategy for choosing the regularization parameter α, which is also known as Lepskij or balancing principle, see [1, 13, 14]. This strategy has been well-established in the recent years since it is easy to implement and applicable under relatively weak technical assumptions. For given (sufficiently small) $\alpha_0 > 0$, $q > 1$ and maximal index $M > 0$ we define the (finite) sequence

$$\{\alpha_j := q^j \alpha_0, \,:\, 0 \le j \le M\}. \tag{5.1}$$

The maximal index M is chosen such that $\alpha_M \le a^2 = \|A^*A\|^2$. Then we can present the following a-posteriori choice of the regularization parameter α.

Definition 4 [Lepskij–Principle]. Let the sequence $\{\alpha_j\}$ be defined by (5.1). For given $\nu \ge 0$ and discretization levels $\xi_{h,j} \le C\alpha_j^{\nu}$ and $\eta_{h,j} \le C\alpha_j^{\frac{\nu}{\min\{2\nu,2\}}}$ for some constant $C > 0$ we calculate $x^\delta_{\alpha_j,h} := R_{\alpha_j,h} y^\delta$. We choose the regularization parameter $\alpha_L := \alpha_{j_L}$ such that

$$j_L := \max\left\{ j \le M \,:\, \|x^\delta_{\alpha_i,h} - x^\delta_{\alpha_j,h}\| \le 4\,C_1 \frac{\delta}{\sqrt{\alpha_i}},\ \forall\, i \le j \right\}.$$

Then $x^\delta_{\alpha,h} := x^\delta_{\alpha_L,h}$ is chosen as regularized solution of (1.1).

We summarize the most important facts. The main idea of the balancing principle is based on the decomposition of the approximation error of regularized solutions into two parts which both depend on the regularization parameter α. We state the assumption in detail below.

Presumption 1. For each $0 < \alpha \le a^2$ and given data y^δ let x^δ_α denote any regularized solution of (1.1) satisfying

$$\|x^\delta_\alpha - x^\dagger\| \le \frac{1}{2}\left(\psi(\alpha) + \phi(\alpha)\right)$$

for a known non-increasing function $\psi(\alpha)$, which can depend on δ and an unknown non-decreasing (index) function $\phi(\alpha)$.

Here we have $\psi(\alpha) = 2C_1 \dfrac{\delta}{\sqrt{\alpha}}$. Now we can establish main theoretical results of the balancing principle, see [14].

Proposition 4. *Let the regularization parameter $\alpha_L > 0$ be chosen by (5.1). Assume $\alpha_0 > 0$ and j_{max} are chosen such that $\phi(\alpha_0) < \psi(\alpha_0)$ and $\alpha_0 < \alpha_L < \alpha_{max}$. If Presumption 1 is valid for some function $\phi(\alpha)$, $\alpha > 0$, then the estimate*

$$\|x^\delta_{\alpha_L} - x^\dagger\| \le 3\sqrt{D}\min\left\{\psi(\alpha_j) + \phi(\alpha_j), 0 \le j \le j_{\max}\right\}$$

holds true. Moreover,

(i) If $x^\dagger \in \mathcal{R}\left((A^*A)^\nu\right)$*, i.e.* $x^\dagger = (A^*A)^\nu \omega$, $\omega \in \mathcal{X}$ *with* $\|\omega\| \le R$*, then*

$$\|x^\delta_{\alpha_L,h} - x^\dagger\| \le 6\sqrt{q}\max\{C_1, C_3 R\}\delta^{\frac{2\nu}{2\nu+1}};$$

(ii) If $x^\dagger \notin \mathcal{R}\left((A^*A)^\nu\right)$ *with distance function* $d(R) := d_{t^\nu}(R; x^\dagger)$ *then*

$$\|x^\delta_{\alpha_L,h} - x^\dagger\| \le 6\sqrt{q}\max\{C_1, C_3 + C_4\} d\left(\Phi^{-1}(\delta)\right).$$

The first estimate is an immediate application of [14, Corollary 1]. The proof of the convergence rates is similar to the proof of [6, Proposition 6.2].

6 A Numerical Example

For illustration of the above considerations we present a short numerical example. For simplicity we choose the linear operator $A : L^2(0,1) \longrightarrow L^2(0,1)$ from Example 1. As well-known, for the adjoint operator $A^* : L^2(0,1) \longrightarrow L^2(0,1)$

$$(A^* y)(t) := \int_t^1 y(\tau)\, d\tau, \quad t \in [0,1], \;\; y \in L^2(0,1),$$

holds. The simple structure of the operator allows us to quote specific source conditions explicitly. In particular,

$$x \in \mathcal{R}\left((A^*A)^{\frac{1}{2}}\right) \;\Leftrightarrow\; x \in \mathcal{R}(A^*) \;\Leftrightarrow\; x' \in L^2(0,1) \;\text{ and }\; x(1) = 0$$

and

$$x \in \mathcal{R}(A^*A) \;\Leftrightarrow\; x'' \in L^2(0,1),\; x'(0) = 0 \;\text{ and }\; x(1) = 0$$

hold. As projections we choose $P_h = Q_h : L^2(0,1) \longrightarrow L^2(0,1)$ with

$$x \in L^2(0,1), \quad P_h x := \sum_{j=1}^n x_j \chi_j, \quad x_j := n \int_{t_{j-1}}^{t_j} x(\tau)\, d\tau,\; j = 1, 2, \ldots, n,$$

and

$$\chi_j(t) := \begin{cases} 1, & t \in (t_{j-1}, t_j), \\ 0, & \text{else}, \end{cases} \qquad t_j := \frac{j}{n}, \;\; j = 0, 1, \ldots, n$$

defines the approximation by piece-wise constant functions. We set $h := 1/n$. With $y := A\,x$ we have for the approximation error η_h

$$\|(I - P_h)A\,x\|_{L^2} = \|(I - P_h)y\|_{L^2} \le h\,\|y'\|_{L^2} = h\,\|x\|_{L^2},$$

see e.g. [21, Theorem 6.1]. Hence we can set $\eta_h := h$. Finding a bound for the second approximation error ξ_h we set $AP_h x =: \tilde{y}_h$, where $\tilde{y}_h$ is the approximation of $y = A\,x$ by piece-wise linear functions. Applying known estimates, see [5, Chapter 8], we can present two bounds for the approximation error depending on the smoothness of x. Without further assumptions we have

$$\|A(I - P_h)x\|_{L^2} = \|y - \tilde{y}_h\|_{L^2} \le C\,h\,\|y'\|_{L^2} = C\,h\,\|x\|_{L^2}$$

for some constant $C > 0$. If, in addition, $x' \in L^2(0,1)$ we obtain a better estimate

$$\|y - \tilde{y}_h\|_{L^2} \leq C\, h^2 \|y''\|_{L^2} = C\, h^2 \|x'\|_{L^2}.$$

In general, ξ_h is given by the first estimate. However, a careful study of the above proofs indicates, that we can also make use of the second estimate supposing sufficient smoothness of the underlying elements. This will be done here in the specific choice of the discretization levels in the numerical example.

As regularization method we apply Tikhonov's regularization, i.e. $g_\alpha(t) := \dfrac{1}{\alpha + t}$, $t \geq 0$. Moreover, the estimate (2.1) in Definition 1 holds with $C_1 = \frac{1}{2}$. Since the regularization method has maximal qualification $\varrho(t) = t$, $t \geq 0$, it is natural to choose $\nu := 1$. Hence, the choice $n \sim 1/\sqrt{\alpha}$ satisfies the required asymptotic behavior of the discretization errors ξ_h and η_h by taking into account the additional smoothness of the elements $y = A\,x$.

For examining convergence rates we consider the following three sample functions:

$$x_1^\dagger(t) := (t - 0.5)^2, \quad x_2^\dagger(t) := t(1-t), \quad x_3^\dagger(t) := \frac{1 - t^2}{4}, \quad t \in [0,1].$$

They are chosen such that $x_3^\dagger \in \mathcal{R}(A^*A)$, $x_2^\dagger \in \mathcal{R}(A^*)$ but $x_2^\dagger \notin \mathcal{R}(A^*A)$ and $x_1^\dagger \notin \mathcal{R}(A^*)$. Moreover, based on example 1 we can state the source conditions more precisely. Since $x_1^\dagger(t)$ is differentiable we have $x_1^\dagger \in \mathcal{R}((A^*A)^\mu)$ for all $0 < \mu < \frac{1}{4}$ and associated distance function $d_\varphi(R; x_1^\dagger) \leq K\, R^{-\frac{1}{3}}$, $R > 0$, for some constant $K > 0$ and $\varphi(t) = t$, $t \geq 0$. On the other hand, since $x_2^\dagger(t)$ is twice differentiable we can conclude $\frac{d}{dt} x_2^\dagger \in \mathcal{R}((A^*A)^\mu)$ and hence $x_2^\dagger \in \mathcal{R}((A^*A)^{\mu + \frac{1}{2}})$ for all $0 < \mu < \frac{1}{4}$. This implies a distance function $d_\varphi(R; x_2^\dagger) \leq \tilde{K}\, R^{-3}$, $R > 0$, for another constant $\tilde{K} > 0$. Furthermore, we choose

$$M := 31, \quad \alpha_{31} := 1, \quad q := \sqrt{2}, \quad n_0 = 20,$$

where n_0 denotes the number of discretizations steps at the coarsest discretization level. The maximal number n of unknowns is given by $n_{max} = 5120$. The noisy data is generated as follows: for $n = n_{max} = 5120$ let e be piece-wise linear function with Gaussian variables $e(t_j) \sim N(0,1)$, $1 \leq j \leq n_{max}$, and $e(0) = 0$ on the finest grid. Then we set

$$y^\delta := y + \frac{e}{\|e\|_{L^2}} \delta, \quad \delta \geq 0.$$

In order to access only function values on the finest grid we suggest the following simplified coarsening strategy: the grid is coarsened always after 4 steps by halving n. Since $\alpha_{j+4} = q^4 \alpha_j = 4\alpha_j$, the asymptotic behavior of the discretization level remains correct.

The numerical results are presented in Table 1 and 2. For the functions $x_1^\dagger(t)$ and $x_2^\dagger(t)$ Table 1 contains the regularization parameter α_L obtained by the balancing principle as well as the discretization levels $n = 1/h$ and the approximation error $\|x_{\alpha_L,h}^\delta - x_l^\dagger\|_{L^2}$, $l = 1, 2, 3$, for different noise levels δ.

Table 1. Approximation errors for the functions $x_1^\dagger$ and $x_2^\dagger$ depending on δ.

	$x_1^\dagger$			$x_2^\dagger$		
δ	α_L	n	$\Vert x^\delta_{\alpha_L,h} - x_1^\dagger\Vert$	α_L	n	$\Vert x^\delta_{\alpha_L,h} - x_2^\dagger\Vert$
10^{-2}	$2.50 \cdot 10^{-1}$	20	$8.89 \cdot 10^{-2}$	$6.25 \cdot 10^{-2}$	80	$5.57 \cdot 10^{-2}$
10^{-3}	$2.76 \cdot 10^{-3}$	160	$3.93 \cdot 10^{-2}$	$5.52 \cdot 10^{-3}$	160	$1.35 \cdot 10^{-2}$
10^{-4}	$1.22 \cdot 10^{-4}$	640	$1.74 \cdot 10^{-2}$	$9.77 \cdot 10^{-2}$	640	$3.87 \cdot 10^{-3}$
10^{-5}	$5.39 \cdot 10^{-6}$	2560	$7.88 \cdot 10^{-3}$	$1.73 \cdot 10^{-4}$	1280	$1.07 \cdot 10^{-3}$
10^{-6}	$2.38 \cdot 10^{-7}$	5120	$3.24 \cdot 10^{-3}$	$3.05 \cdot 10^{-5}$	2560	$2.99 \cdot 10^{-4}$

Table 2. Approximation errors for the function $x_3^\dagger$ depending on δ.

	$x_3^\dagger$					
δ	α_L	n	$\Vert x^\delta_{\alpha_L,h} - x_3^\dagger\Vert$	α_{opt}	n	$\Vert x^\delta_{\alpha_{opt},h} - x_3^\dagger\Vert$
10^{-2}	$8.84 \cdot 10^{-2}$	40	$3.19 \cdot 10^{-2}$	$1.56 \cdot 10^{-2}$	160	$1.58 \cdot 10^{-2}$
10^{-3}	$1.56 \cdot 10^{-2}$	160	$7.21 \cdot 10^{-3}$	$3.91 \cdot 10^{-3}$	320	$2.56 \cdot 10^{-3}$
10^{-4}	$3.91 \cdot 10^{-3}$	320	$1.86 \cdot 10^{-3}$	$6.91 \cdot 10^{-4}$	640	$5.32 \cdot 10^{-4}$
10^{-5}	$6.91 \cdot 10^{-4}$	640	$3.61 \cdot 10^{-4}$	$1.73 \cdot 10^{-4}$	1280	$1.39 \cdot 10^{-4}$
10^{-6}	$8.63 \cdot 10^{-5}$	1280	$7.99 \cdot 10^{-5}$	$3.05 \cdot 10^{-5}$	2560	$4.31 \cdot 10^{-5}$

Here, δ in fact describes the absolute noise level. The lower source condition on $x_1^\dagger(t)$ implies a faster decay of the regularization parameter α_L as $\delta \to 0$ than in the case of $x_2^\dagger(t)$. Since the discretization level h is connected to the choice of the regularization parameter we observe the following: starting with a coarser discretization for $x_1^\dagger(t)$ in comparison to $x_2^\dagger(t)$ for $\delta = 10^{-2}$ we end up at a finer discretization for $x_1^\dagger(t)$ when we have $\delta = 10^{-6}$ since the chosen regularization parameter α_L is much smaller than in the case of $x_2^\dagger$. The same data can be found in Table 2 for the function $x_3^\dagger(t)$. Here, additionally the 'optimal' regularization is presented, i.e. the regularization parameter $\alpha = \alpha_{opt}$ satisfies

$$\Vert x^\delta_{\alpha_{opt},h} - x_l^\dagger\Vert_{L^2} = \min\left\{\Vert x^\delta_{\alpha_j,h} - x_l^\dagger\Vert_{L^2} \ : \ 0 \le j \le M\right\}, \quad l = 1,2,3.$$

As we can observe, the approximation error of the best regularization parameter α_{opt} is amplified with a factor 2–3 by the balancing principle. This is a better approximation than the theoretical result of Proposition 4, where the factor 6 is deduced. We also remark that the regularization parameter α_L is chosen somewhat too large in comparison to the optimal value α_{opt} which is a known property of the balancing principle. This again leads to the coarser discretization levels for α_L compared to α_{opt}.

The convergence rates of the approximation error depending on the noise level δ are presented in Figure 1 in log-log-diagrams for all the sample functions. The solid line shows the approximation of the error by a function of power-type, i.e. $\Vert x^\delta_{\alpha_L,h} - x_l^\dagger\Vert_{L^2} \sim \delta^{\mu_l}$ for some $\mu_l > 0$, $l = 1,2,3$. Using the described (approximate) source conditions we expect the rates $\mu_1 = \frac{1}{3}$, $\mu_2 = \frac{3}{5} = 0.6$

ht

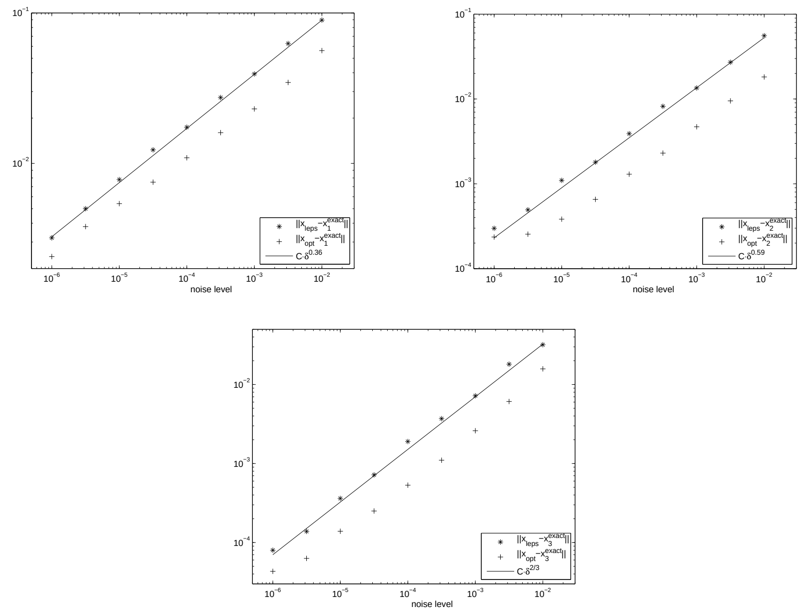

Figure 1. Approximation errors depending on the noise level δ for $x_1^\dagger$ (top) $x_2^\dagger$ (middle) and $x_3^\dagger$ (bottom).

and $\mu_3 = \frac{2}{3}$. Comparing this with the numerical result we observe that the numerical convergence rates rather well coincide with the theoretical values.

References

[1] F. Bauer and T. Hohage. A Lepskij-type stopping rule for regularized Newton methods. *Inverse Problems*, **21**:1979–1991, 2005. (doi:10.1088/0266-5611/21/6/011)

[2] J. Baumeister. *Stable solution of Inverse problems*. Vieweg Braunschweig, 1987.

[3] D. Düvelmeyer, B. Hofmann and M. Yamamoto. Range inclusions and approximative source conditions with general benchmark functions. *Num. Func. Anal. Optim.*, **28**(11):1245–1261, 2007. (doi:10.1080/01630560701749649)

[4] H.W. Engl, M. Hanke and A. Neubauer. *Regularization of Inverse Problems*. Kluwer Academic Publishers, Dortrecht, 1996.

[5] W. Hackbusch. *Elliptic Differential equations – Theory and numerical Treatment*. Springer-Verlag Berlin, 1992.

[6] T. Hein. Regularization in Banach spaces – convergence rates by approximative source conditions. *J. Inv. Ill-Posed Problems*, **17**:27–42, 2009. (doi:10.1515/JIIP.2009.004)

[7] B. Hofmann. Approximative source conditions in Tikhonov–Phillips regularization and consequences for inverse problems with multiplication operators. *Mathematical Methods in Applied Sciences*, **29**:351–371, 2006. (doi:10.1002/mma.686)

[8] B. Hofmann, D. Düvelmeyer and K. Krumbiegel. Approximative source conditions in Tikhonov regularization – new analytical results and some numerical studies. *Math. Model. Anal.*, **10**:41–56, 2006.

[9] B. Hofmann and P. Mathé. Analysis of profile functions for general linear regularization methods. *SIAM J.Numer.Anal.*, **45**(3):1122–1141, 2007. (doi:10.1137/060654530)

[10] B. Hofmann and P. Mathé. How general are general source conditions? *Inverse Problems*, **24**:015009, 2008. (doi:10.1088/0266-5611/24/1/015009)

[11] T. Hohage. Logarithmic convergence rates of the iteratively regularized Gauss-Newton method for an inverse potential and an inverse scattering problem. *Inverse Problems*, **13**(5):1279–1299, 1997. (doi:10.1088/0266-5611/13/5/012)

[12] B. Kaltenbacher. Regularization by projection with a posteriori discretization level choice for linear and nonlinear ill-posed problems. *Inverse Problems*, **16**:1523–1539, 2000. (doi:10.1088/0266-5611/16/5/322)

[13] O.V. Lepskij. A problem of adaptive estimation in Gaussian withe noise. *Teor. Veroyatn.Primen.*, **35**:459–470, 1990.

[14] P. Mathé. The Lepskij principle revisited. *Inverse Problems*, **22**:L11–L15, 2006. (doi:10.1088/0266-5611/22/3/L02)

[15] P. Mathé and S.V. Pereverzev. Discretization strategies for linear ill-posed problems in variable Hilbert scales. *Inverse Problems*, **19**:1263–1277, 2003. (doi:10.1088/0266-5611/19/6/003)

[16] P. Mathé and S.V. Pereverzev. Geometry of linear ill-posed problems in variable Hilbert scales. *Inverse Problems*, **19**:789–803, 2003. (doi:10.1088/0266-5611/19/3/319)

[17] M.T. Nair, E. Schock and U. Tautenhahn. Morozov's discrepancy principle under general source conditions. *Zeitschrift für Analysis u. Anwendungen*, **299**:199–214, 2003.

[18] F. Natterer. Regularisierung schlecht gestellter Probleme durch Projektionsverfahren. *Numer.Math.*, **28**:329–341, 1977. (doi:10.1007/BF01389972)

[19] A. Neubauer. On converse and saturation results for regularization methods. In E. Schock(Ed.), *Beiträge zur Angewandten Mathematik und Informatik, Helmut Brakhage zu Ehren*, pp. 262–270. Shaker, Aachen, 1994.

[20] R. Plato and G. Vainniko. On the regularization of projection methods for solving ill-posed problems. *Numer.Math.*, **57**:63–79, 1990. (doi:10.1007/BF01386397)

[21] L.L. Schumaker. *Spline Functions: Basic Theory.* Wiley New York, 1991.

[22] U. Tautenhahn. Optimality for ill-posed problems under general source conditions. *Numer.Funct.Anal.Optim.*, **19**(3):377–398, 1998. (doi:10.1080/01630569808816834)

[23] E. Zeidler. *Nonlinear Functional Analysis and its Applications III.* Springer-Verlag New York, 1985.