Semiiterative Regularization in Hilbert Scales*

Herbert Egger

Spezialforschungsbereich F013, Johann Radon Institute
for Computational and Applied Mathematics, A-4040 Linz, Austria
email: herbert.egger@oeaw.ac.at

Abstract

In this paper we investigate the regularization properties of semiiterative regularization methods in Hilbert scales for linear ill-posed problems and perturbed data.

It is well known that Landweber iteration can be remarkably accelerated by polynomial acceleration methods leading to the notion of *optimal speed of convergence*, which can be obtained by several efficient two-step methods, e.g., the $\nu-$methods by Brakhage. It was observed in [3] that a similar speed of convergence, i.e., similar iteration numbers yielding optimal convergence rates, can be obtained if Landweber iteration is performed in Hilbert scales.

Combining both ideas, we show that semiiterative methods can be further accelerated yielding optimal convergence rates with only the square root of iterations compared to semiiterative regularization methods or Landweber iteration in Hilbert scales. We conclude with several examples and numerical tests confirming the theoretical results, including a comparison to the method of conjugate gradients.

1 Introduction

In this paper, we study inverse problems of the form

$$Tx = y, \qquad y \in \mathcal{R}(K), \tag{1.1}$$

where $T : \mathcal{X} \to \mathcal{Y}$ is a linear bounded operator between infinite dimensional Hilbert spaces $\mathcal{X}$ and $\mathcal{Y}$ with range $\mathcal{R}(T) \subset \mathcal{Y}$. It is well known (see, e.g., [6]) that, if $\mathcal{R}(T)$ is not closed, the Moore-Penrose inverse $T^\dagger$ defined on $\mathcal{D}(T^\dagger) = \mathcal{R}(T) + \mathcal{R}(T)^\perp$, is unbounded and the solution of (1.1) is ill-posed; i.e. a solution of (1.1) does not depend continuously on the right hand side, thus it has to be regularized.

Especially for large scale problems, iterative regularization algorithms have turned out to be an attractive alternative to Tikhonov regularization, which is probably the

*supported by the Austrian Science Fund "Fonds zur Förderung der wissenschaftlichen Forschung (FWF)" under grant SFB F013, "Numerical and Symbolic Scientific Computing", Project F1308

most well known regularization method (see, e.g., [4, 7]). Application of Landweber iteration (cf. [12]),

$$x_k = x_{k-1} + \omega T^*(y - Tx_{k-1}), \qquad k \geq 1, \tag{1.2}$$

with $0 < \omega < 2/\|T^*T\|$ to the solution of inverse problems has been investigated intensively in the literature (see, e.g., [1, 4, 9] and the references cited there). Note that in finite dimensions, (1.2) corresponds to Richardson iteration (successive approximation) applied to the normal equation

$$T^*Tx = T^*y. \tag{1.3}$$

If $y \in \mathcal{R}(T)$, then the iterates x_k converge to $T^\dagger y$; if however only perturbed data y^δ with a known upper bound on the noise level

$$\|y - y^\delta\| \leq \delta \tag{1.4}$$

are known and $y^\delta \notin \mathcal{R}(T)$, which is most probable if (1.1) is ill-posed and $\mathcal{R}(T)$ is not closed, $\|x_k\|$ tends to infinity.

Iterative methods are turned into regularization algorithms by stopping the iteration after an adequate number k_* of steps. Besides *a priori* stopping rules, which require knowledge of the smoothness of $x^\dagger - x_0$ in terms of spaces $\mathcal{R}((T^*T)^\mu)$, the discrepancy principle (cf. [4, 14])

$$\|y^\delta - Tx_{k_*}\| \leq \tau\delta < \|y^\delta - Tx_k\|, \qquad 0 \leq k < k_*, \tag{1.5}$$

with $\tau > 1$ has turned out to be an appropriate *a posteriori* stopping rule yielding optimal convergence rates for Landweber iteration for linear problems, i.e., if

$$x^\dagger - x_0 \in \mathcal{R}((T^*T)^\mu), \qquad \mu > 0 \tag{1.6}$$

then

$$\|x^\delta_{k_*} - x^\dagger\| = o(\delta^{\frac{2\mu}{2\mu+1}}) \quad \text{and} \quad k_* \sim \delta^{-\frac{2}{2\mu+1}}, \tag{1.7}$$

see, e.g., [4].

The main drawback of Landweber iteration is the large number of iterations needed to obtain the optimal convergence rates, cf. (1.7). To speed up the method, several *semiiterative methods* (polynomial acceleration methods) have been investigated (see, e.g., [8] for an overview). In our numerical experiments, we will use the $\nu-$methods proposed by Brakhage [2], for which the number of iteration can be bounded by

$$k_* \sim \delta^{-\frac{1}{2\mu+1}} \quad \text{for} \quad 0 < \mu \leq \nu - 1/2 \tag{1.8}$$

in case of stopping with the discrepancy principle (1.5) (see Theorem 2.1). Thus only the square root of the number of iterations as compared to Landweber iteration have to be performed to get optimal convergence rates. However, the $\nu-$methods show a saturation phenomenon, which was not present for Landweber iteration, i.e., the optimal rates and (1.8) hold only for $\mu \leq \nu$ respectively $\mu \leq \nu - 1/2$ if the iteration is stopped according to (1.5).

Regularization in Hilbert scales was introduced by Natterer [15] in the framework of Tikhonov regularization for linear problems and it has been investigated for more general regularization methods in [4, 20] by means of spectral theory. Originally, Hilbert scale regularization was introduced to increase the range of optimal convergence for Tikhonov regularization [15, 18] and Landweber iteration for nonlinear inverse problems [19]. In [3], the case $s \leq 0$ (undersmoothing) was focused, and it was shown that the application of Hilbert scales can be understood as preconditioning in this case, i.e., the number of iterations needed to get optimal convergence can essentially be reduced to (1.8). Additionally, the results were derived under more general than the usual assumptions for regularization in Hilbert scales (cf. Section 3).

The aim of this paper is to show that polynomial acceleration methods in combination with the Hilbert scale approach can lead to further acceleration of iterative methods yielding optimal rates of convergence with the stopping index bounded by $k_* \sim \delta^{-\frac{1}{2(2\mu+1)}}$. In case of mildly ill-posed problems, i.e., if the singular values σ_n of the operator T decay like $O(n^{-\alpha})$ for some $0 < \alpha < 1$, this bound is even better than the one for the conjugate gradient method $k_* \leq \delta^{-\frac{1}{(2\mu+1)(1+\alpha)}}$, cf. [4, Theorem 7.14]. The faster convergence of a Hilbert scale $\nu-$method is illustrated in Example 5.1 and 5.3.

The paper is organized as follows: In Sections 2 and 3, we shortly repeat the main results on convergence of semiiterative regularization methods and regularization in Hilbert scales. The convergence analysis for semiiterative methods in Hilbert scales is given in Section 4. We conclude with numerical examples comparing the proposed method with standard Landweber iteration and $\nu-$methods, Landweber iteration in Hilbert scales and the conjugate gradient method.

2 Accelerated Landweber methods

While for Landweber iteration (1.2) only information on the last iterate x_{k-1}^δ is used to construct the new approximation x_k^δ, semiiterative methods make use of the all the approximations for $T^\dagger y$ obtained so far: A basic step of a semiiterative method has the form

$$\begin{aligned} x_k^\delta = \mu_{1,k} x_{k-1}^\delta + \ldots + \mu_{k,k} x_0 + \omega_k T^*(y^\delta - T x_k^\delta), \quad k \geq 1 \\ \textstyle\sum_{i=1}^k \mu_{i,k} = 1, \qquad \omega_k \neq 0. \end{aligned} \tag{2.1}$$

By x_k we denote the iterates obtained with y^δ in (2.1) replaced by the exact data y. Obviously $x_k^\delta - x_0 \in \mathcal{K}_k(T^*T, T^*(y^\delta - Tx_0))$, where

$$\mathcal{K}_k(T^*T, p) := [p, T^*Tp, \ldots, (T^*T)^{k-1}p] \tag{2.2}$$

denotes the kth Krylov subspace of T^*T with respect to p. Consequently, there exist polynomials $g_k(\lambda)$ and $r_k(\lambda) := 1 - \lambda g_k(\lambda)$ of degree $(k-1)$ respectively k, such that

$$x_k - x^\dagger = r_k(T^*T)(x_0 - x^\dagger) \qquad \text{and} \qquad x_k^\delta - x_k = g_k(T^*T)T^*(y^\delta - y). \tag{2.3}$$

In other words, the approximation error $x_k - x^\dagger$ is determined by the *residual polynomials* r_k, while the propagated data error $x_k^\delta - x_k$ is determined by the *iteration polynomials* g_k. A semiiterative method (2.1) is said to have *optimal speed of convergence* (cf. [8]) for $x^\dagger - x_0 \in \mathcal{R}((T^*T)^\mu)$, if $\|x_k^\delta - x^\dagger\| = O(k^{-2\mu})$. This can be guaranteed, if

$$\|\lambda^\mu r_k(\lambda)\|_{\mathcal{C}[0,1]} = \omega_\mu(k) = O(k^{-2\mu}), \tag{2.4}$$

see [4, Section 6.2]. The following theorem guarantees optimal convergence of semiiterative regularization methods:

Theorem 2.1 *(See [4, Theorem 6.11]) Let $y \in \mathcal{R}(T)$, and let the residual polynomials r_k satisfy* (2.4) *for some $\mu_0 > 0$. Then the semiiterative method* (2.1) *is a regularization method of optimal order for $T^\dagger y \in \mathcal{R}((T^*T)^\mu)$ with $0 < \mu \leq \mu_0 - 1/2$ provided the iteration is stopped with $k_* = k_*(\delta, y^\delta)$ according to the discrepancy principle* (1.5) *with fixed $\tau > \sup_{k \in \mathcal{N}} \|r_k\|_{\mathcal{C}[0,1]}$. In this case we have $k_* = O(\delta^{-\frac{1}{2\mu+1}})$ and $\|x_k^\delta - x^\dagger\| = O(\delta^{\frac{2\mu}{2\mu+1}})$.*

Note that also $o(\cdot)$ can be derived for the error $\|x_k^\delta - x^\dagger\|$ (see [4]). Of particular importance is the case when the residual polynomials $\{r_k\}$ are an orthogonal sequence with respect to some weight function. In this case, the residual polynomials satisfy a three-term-recurrence, which also carries over to the iterates, i.e.,

$$x_k^\delta = x_{k-1}^\delta + \mu_k(x_{k-1}^\delta - x_{k-2}^\delta) + \omega_k T^*(y^\delta - T x_{k-1}^\delta), \quad k \geq 1, \tag{2.5}$$

with $x_{-1}^\delta = 0$. A specific choice of such orthogonal polynomials defines the $\nu-$methods by Brakhage via $\mu_1 = 0$, $\omega_1 = (4\nu+2)/(4\nu+1)$ and

$$\begin{aligned} \mu_k &= \tfrac{(k-1)(2k-3)(2k+2\nu-1)}{(k+2\nu-1)(2k+4\nu-1)(2k+2\nu-3)}, \\ \omega_k &= 4\tfrac{(2k+2\nu-1)(k+\nu-1)}{(k+2\nu-1)(2k+2\nu-1)}, \qquad k > 1, \end{aligned} \tag{2.6}$$

which have optimal speed of convergence for $x^\dagger - x_0 \in X_\mu$, $0 \leq \mu \leq \nu$, i.e., (2.4) holds with $\mu = \nu$.

Remark 2.2 The notion *optimal speed of convergence* is explained by the fact that the minimal modulus of convergence $\omega_\nu^*(k) = O(k^{-2\nu})$ (see Brakhage [2]), thus no faster faster convergence than $O(k^{-2\mu})$ can be expected for the approximation error for semiiterative methods in general.

The $1/2-$method of Brakhage corresponds to the Chebychev method of Nemirovskii and Polyak, investigated somewhat earlier in [17].

Since the $\nu-$methods have a finite qualification ν, i.e., they satisfy (2.4) only for $\mu \leq \nu$, it is not surprising that the discrepancy principle guarantees optimal convergence rates only for $0 < \mu \leq \nu - 1/2$. In [4, Section 6], an improved a posteriori stopping rule is investigated yielding optimal convergence for $0 < \mu \leq \nu$. There also the relation of the $\nu-$methods with iterated Tikhonov regularization is discussed.

3 Regularization in Hilbert scales

Before we recall some results on regularization in Hilbert scales, we shortly repeat the definition of a Hilbert scale (see [11]):

Let L be a densly defined unbounded selfadjoint strictly positive operator in $\mathcal{X}$. Then $(\mathcal{X}_s)_{s\in\mathcal{R}}$ denotes the Hilbert scale induced by L if $\mathcal{X}_s$ is the completion of $\bigcap_{k=0}^{\infty} D(L^k)$ with respect to the Hilbert space norm $\|x\|_s := \|L^s x\|_{\mathcal{X}}$; obviously $\|x\|_0 = \|x\|_{\mathcal{X}}$ (see [11] or [4, Section 8.4] for details).

Regularization in Hilbert scales was introduced by Natterer [15] in order to improve convergence rates for Tikhonov regularization. In [19], Landweber iteration for nonlinear problems, which exhibits similar saturation phenomena as Tikhonov regularization (i.e., optimal convergence only for $x^\dagger - x_0 \in \mathcal{R}((T^*T)^\mu)$, $\mu \leq 1/2$) has been shifted to Hilbert scales (with $s > 0$) in order to get rid of the restriction $\mu \leq 1/2$.

In [3], the application of Hilbert scales to iterative regularization methods has been investigated from a different point of view. There, the emphasis has been put on the case $s \leq 0$, in which case the Hilbert scale operator L^{-2s} appearing in the modified Landweber iteration

$$x_{k+1}^\delta = x_k^\delta + L^{-2s}T^*(y^\delta - Tx_k^\delta), \qquad k \geq 0 \tag{3.1}$$

acts as a preconditioner for the adjoint operator T^*. As a consequence, the operator M_s appearing in the preconditioned normal equation

$$M_s x := L^{-2s}T^*Tx = L^{-2s}T^*y^\delta \tag{3.2}$$

has a smaller degree of ill-posedness than T^*T, while still being self-adjoint in $\mathcal{X}_s$. For a finite dimensional approximation, this means that the condition number of the operator M_s is of the same order as the condition number of T, and substantially smaller than the one of T^*T appearing in the normal equations. This yields a smaller stopping index determined by the discrepancy principle (1.5).

We shortly recall the main assumptions and convergence results for Landweber iteration for linear problems in Hilbert scales (cf. [3]):

Assumption 3.1

(A1) $Tx = y$ *has a solution* $x^\dagger$.

(A2) $\|Tx\| \leq \overline{m}\|x\|_{-a}$ *for all* $x \in \mathcal{X}$ *and some* $a > 0, \overline{m} > 0$. *Moreover, the extension of* T *to* $\mathcal{X}_{-a}$ *(again denoted by* T*) is injective.*

(A3) $B := TL^{-s}$ *is such that* $\|B\|_{\mathcal{X},\mathcal{Y}} \leq 1$, *where* $s \geq -a$.

Usually, for the analysis of regularization methods in Hilbert scales, a stronger condition than (A2) is used, namely (cf, e.g., [15, 18])

$$\|Tx\| \sim \|x\|_{-a} \qquad \text{for all} \quad x \in \mathcal{X}, \tag{3.3}$$

where the number a can be interpreted as the *degree of ill-posedness*. However, if $s \leq 0$, an estimate from below (possibly in a weaker norm), e.g.,

$$\|Tx\| \geq \underline{m}\|x\|_{-\tilde{a}} \quad \text{for all} \quad x \in \mathcal{X} \quad \text{and some} \quad \tilde{a} \geq a,\ \underline{m} > 0, \tag{3.4}$$

is only needed to interpret the natural source condition $x^\dagger - x_0 \in \mathcal{X}_u^s$, i.e.,

$$x^\dagger - x_0 = L^{-s}(B^*B)^{\frac{u-s}{2(a+s)}} w, \quad \text{for some} \quad w \in \mathcal{X} \tag{3.5}$$

in terms of the Hilbert scale $\{\mathcal{X}_s\}_{s\in\mathcal{R}}$. The following proposition taken from [3] draws some conclusions from Assumption 3.1.

Proposition 3.2 *Let Assumption 3.1 hold. Then Condition (A2) is equivalent to*

$$\mathcal{R}(T^*) \subset \mathcal{X}_a \qquad \textit{and} \qquad \|T^*w\|_a \le \overline{m}\|w\| \quad \textit{for all} \quad w \in \mathcal{Y}\,. \tag{3.6}$$

Moreover for all $\nu \in [0,1]$ *it holds that* $\mathcal{D}((B^*B)^{-\frac{\nu}{2}}) = \mathcal{R}((B^*B)^{\frac{\nu}{2}}) \subset \mathcal{X}_{\nu(a+s)}$ *and*

$$\|(B^*B)^{\frac{\nu}{2}}x\| \;\le\; \overline{m}^\nu \|x\|_{-\nu(a+s)} \qquad \textit{for all} \quad x \in \mathcal{X} \tag{3.7}$$

$$\|(B^*B)^{-\frac{\nu}{2}}x\| \;\ge\; \overline{m}^{-\nu} \|x\|_{\nu(a+s)} \qquad \textit{for all} \quad x \in \mathcal{D}((B^*B)^{-\frac{\nu}{2}}) \tag{3.8}$$

Furthermore, (3.4) *is equivalent to*

$$\begin{aligned} \mathcal{X}_{\tilde a} \subset \mathcal{R}(T^*) \quad \textit{and} \quad & \|T^*w\|_{\tilde a} \ge \underline{m}\|w\| \\ & \textit{for all} \quad w \in \mathcal{N}(T^*)^\perp \ \textit{with}\ T^*w \in \mathcal{X}_{\tilde a} \end{aligned} \tag{3.9}$$

and if (3.4) *holds, then it follows for all* $\nu \in [0,1]$ *that* $\mathcal{X}_{\nu(\tilde a+s)} \subset \mathcal{R}((B^*B)^{\frac{\nu}{2}}) = \mathcal{D}((B^*B)^{-\frac{\nu}{2}})$ *and*

$$\|(B^*B)^{\frac{\nu}{2}}x\| \;\ge\; \underline{m}^\nu \|x\|_{-\nu(\tilde a+s)} \qquad \textit{for all} \quad x \in \mathcal{X} \tag{3.10}$$

$$\|(B^*B)^{-\frac{\nu}{2}}x\| \;\le\; \underline{m}^{-\nu} \|x\|_{\nu(\tilde a+s)} \qquad \textit{for all} \quad x \in \mathcal{X}_{\nu(\tilde a+s)}. \tag{3.11}$$

In our convergence analysis the following shifted Hilbert scale will play an important role:

Definition 3.3 *Let a, s and B be as in Assumption 3.1. We define the shifted Hilbert scale* $\{\mathcal{X}_r^s\}_{r\in\mathcal{R}}$ *by*

$$\begin{aligned} &\mathcal{X}_r^s := \mathcal{D}((B^*B)^{\frac{s-r}{2(a+s)}} L^s) \quad \textit{equipped with the norm} \\ &|\!|\!|x|\!|\!|_r := \|(B^*B)^{\frac{s-r}{2(a+s)}} L^s x\|_{\mathcal{X}}\,. \end{aligned} \tag{3.12}$$

Remark 3.4 Note, that for $s \ne 0$ $\{\mathcal{X}_r^s\}_{r\in\mathcal{R}}$ is no Hilbert scale over $\mathcal{X}$ in general. In particular, $\mathcal{X}_{-r}^s$ is usually not the dual space of $\mathcal{X}_r^s$. Nevertheless, the spaces $\mathcal{X}_r^s$ have some properties (interpolation, embedding), that justify the notion *shifted Hilbert scale* (see [10] for details). If $\tilde T$ denotes the extension of T to $\mathcal{X}_s$ and $\tilde T^*$ denotes the adjoint operator with respect to the spaces $\mathcal{X}_s$ and $\mathcal{Y}$, then $\mathcal{X}_u^s = \mathcal{R}((\tilde T^*\tilde T)^{\frac{u-s}{2(a+s)}})$. Hence, the spaces $\mathcal{X}_u^s$ are natural spaces for sourcewise representations of $x^\dagger - x_0$, if the problem is considered on $\mathcal{X}_s$. We will use this fact several times below.

The following convergence result for Landweber iteration in Hilbert scales is taken from [3]. In the next section we will derive corresponding results also for the class of semiiterative regularization methods.

Proposition 3.5 *Let Assumption 3.1 hold and $-a/2 \le s \le 0$. Additionally, assume $x^\dagger - x_0 \in \mathcal{X}_u^s$ for some $u > 0$. Then*

$$\|x_k^\delta - x^\dagger\| \le c(\delta k^{\frac{a}{2(a+s)}} + k^{-\frac{u}{2(a+s)}} \|x^\dagger - x_0\|_u). \tag{3.13}$$

If the iteration (3.1) *is stopped according to the* a priori *rule* $k^* \sim (\|w\|\delta^{-1})^{\frac{2(a+s)}{a+u}}$ *then*

$$\|x_k^\delta - x^\dagger\| = O(\|w\|^{\frac{a}{a+u}} \delta^{\frac{u}{a+u}}). \tag{3.14}$$

If, alternatively, the iteration is stopped according to the discrepancy principle (1.5) *then*

$$k_* \sim \delta^{-\frac{2(a+s)}{a+u}} \quad \textit{and} \quad \|x_k^\delta - x^\dagger\| = O(\delta^{\frac{u}{a+u}}). \tag{3.15}$$

Remark 3.6 It was mentioned in [3] that, if the usual condition (3.3) holds instead of (A1), then for $0 < u \le a + 2s$ these rates are optimal, i.e., the best possible worst case error bounds under the given source condition. To see that, observe that for $u \le \min(a, a+2s)$ we have $x^\dagger - x_0 \in \mathcal{X}_u^s = \mathcal{R}((T^*T)^\mu)$ with $\mu = \frac{u}{2a}$; hence $\delta^{\frac{u}{a+u}} = \delta^{\frac{2\mu}{2\mu+1}}$.

For $s < 0$, the stopping index of the Hilbert scale method is significantly smaller than the one for Landweber iteration. E.g., the choice $s = -\frac{a}{2}(1-\epsilon)$ and using $u = 2a\mu$ yields approximately the square root of iterations compared to standard Landweber iteration under the weak source condition $x^\dagger - x_0 \in X_\mu$ with $0 < \mu \le \epsilon/2$.

For $u > a+2s$, the source condition (3.5) can in general no longer be interpreted in terms of spaces $\mathcal{R}((T^*T)^\mu)$.

4 Convergence rates for semiiterative regularization methods in Hilbert scales

In this section we investigate the regularization properties of semiiterative methods in Hilbert scales. Note, that under Assumption 3.1, the operator T can be extended to an operator on $\mathcal{X}_s$, and instead of (1.1) one could solve

$$Bz = TL^{-s}z = y^\delta, \qquad x = L^{-s}z. \tag{4.1}$$

Applying polynomial acceleration methods to (4.1) yields

$$z_k - z^\dagger = r_k(B^*B)(z_0 - z^\dagger) \qquad \text{and} \qquad z_k^\delta - z_k = g_k(B^*B)B^*(y^\delta - y), \tag{4.2}$$

with $z^\dagger = L^s x^\dagger$ and $z_0 = L^s x_0$ and consequently

$$\begin{aligned} x_k - x^\dagger &= L^{-s} r_k(B^*B) L^s (x_0 - x^\dagger) \qquad \text{and} \\ x_k^\delta - x_k &= L^{-s} g_k(B^*B)B^*(y^\delta - y), \end{aligned} \tag{4.3}$$

where now the iterates x_k^δ are calculated by the iteration

$$\begin{gathered} x_k^\delta = \mu_{1,k} x_{k-1}^\delta + \ldots + \mu_{k,k} x_0 + \omega_k L^{-2s} T^*(y^\delta - Tx_k^\delta), \quad k \ge 1 \\ \textstyle\sum_{i=1}^k \mu_{i,k} = 1, \qquad \omega_k \ne 0. \end{gathered} \tag{4.4}$$

As for Landweber iteration in Hilbert scales, the residuals $T^*(y^\delta - Tx_k)$ are preconditioned with L^{-2s}.

Remark 4.1 Iterative methods are designed in a way such that the residual polynomials $r_k(\lambda)$ approximate 0, while the iteration polynomials $q_k(\lambda)$ approximate $1/\lambda$. As can be seen from (4.3), we use here the same polynomials as for standard iterative methods. However, the spectrum of the operator B^*B is different from the one of T^*T, in particular, clustering of the eigenvalues at $\lambda = 0$ is somewhat weaker. The application of $L^{-s}(\cdot)L^s$ can also be interpreted as a change of basis.

We are now in the position to state the main results:

Proposition 4.2 *Let Assumption 3.1 hold and $-a/2 \le s \le 0$ and let x_k^δ be defined by the semiiterative method* (4.4) *satisfying* (2.4) *for some $\mu_0 > 0$. Additionally, assume $x^\dagger - x_0 \in \mathcal{X}_u^s$, i.e.,*

$$x^\dagger - x_0 = L^{-s}(B^*B)^{\frac{u-s}{2(a+s)}} w, \tag{4.5}$$

for some $w \in \mathcal{X}$ and $0 < u \le 2(a+s)\mu_0$. Then

$$\|x_k^\delta - x^\dagger\| \le C_u(\delta k^{\frac{a}{(a+s)}} + k^{-\frac{u}{(a+s)}} \|w\|). \tag{4.6}$$

Proof. Using the source condition (4.5) and the representation (4.3), we get with (3.8)

$$\begin{aligned} \|x_k - x^\dagger\| &= \|L^{-s} r_k(B^*B)(B^*B)^{\frac{u-s}{2(a+s)}} w\| \\ &\le c\|(B^*B)^{\frac{u}{2(a+s)}} r_k(B^*B)\| \, \|w\|. \end{aligned}$$

With spectral theory and (2.4) this yields the estimate

$$\|x_k - x^\dagger\| \le c_u k^{-\frac{u}{(a+s)}} \|w\|, \tag{4.7}$$

for the approximation error. Similarly, the propagated data error can be estimated by

$$\begin{aligned} \|x_k^\delta - x_k\| &= \|L^{-s} g_k(B^*B)B^*(y^\delta - y)\| \\ &\le c\delta \, \|(B^*B)^{\frac{a+2s}{2(a+s)}} g_k(B^*B)\|. \end{aligned}$$

Next, we give an estimate for $\|\lambda^\mu g_k(\lambda)\|_{\mathcal{C}[0,1]}$: Using $r_k(\lambda = 1 - \lambda g_k(\lambda))$, we obtain for $0 \le \mu \le 1$

$$\begin{aligned} \lambda^\mu g_k(\lambda) &= \lambda^{\mu-1}(1 - r_k(\lambda)) \\ &= [\lambda^{-1}(1 - r_k(\lambda))]^{1-\mu}[1 - r_k(\lambda)]^\mu. \end{aligned}$$

Now, by the Mean Value Theorem, one can find a $\tilde{\lambda} \in [0,1]$ such that

$$\lambda^{-1}(1 - r_k(\lambda)) = -r_k'(\tilde{\lambda}),$$

which together with Markov's inequality ($|r_k'(\lambda)| \le 2k^2$) and $|r_k(\lambda)| \le 2$ for $\lambda \in [0,1]$ yields

$$\lambda^\mu g_k(\lambda) \le 2k^{2(1-\mu)} \quad \text{for} \quad \lambda \in [0,1].$$

Note that by $-a/2 \leq s \leq 0$ we always have $0 \leq \frac{a+2s}{2(a+s)} \leq 1$ and thus

$$\|x_k^\delta - x_k\| \leq 2c\delta k^{\frac{a}{a+s}}. \tag{4.8}$$

■

In order to get convergence rates in terms of δ it remains to bound the number of iterations k_* in terms of δ. Note that Proposition 4.2 already guarantees convergence, if k_* is chosen such that $\delta k^{\frac{a}{a+s}} \to 0$ with $k \to \infty$. In order to derive optimal rates in terms of δ, we can use the analogy of iterative methods in Hilbert scales with iterations in $\mathcal{X}_s$ (cf. (4.2)).

Theorem 4.3 *Let the assumptions of Proposition 4.2 be satisfied and x_k^δ be generated by the semiiterative method in Hilbert scales* (4.4) *satisfying* (2.4) *for some $\mu_0 > 0$. If the iteration is stopped according to the* a priori *stopping rule $k_* = O(\delta^{\frac{a+s}{a+u}})$ then*

$$\|x_k^\delta - x^\dagger\| = O(\delta^{\frac{u}{a+u}}) \tag{4.9}$$

for $x^\dagger - x_0 \in \mathcal{X}_u^s$ with $0 < u \leq 2(a+s)\mu_0 + s$.

If, alternatively, the iteration is stopped according to the discrepancy principle (1.5)*, then*

$$k_* = O(\delta^{\frac{a+s}{a+u}}) \quad \textit{and} \quad \|x_k^\delta - x^\dagger\| = O(\delta^{\frac{u}{a+u}}) \tag{4.10}$$

for $x^\dagger - x_0 \in \mathcal{X}_u^s$ with $0 < u \leq 2(a+s)(\mu_0 - 1/2) + s$.

Proof. The first result follows immediately with Proposition 4.2. For the second result, observe that (4.4) can be interpreted as semiiterative method for the problem $\tilde{T}x = y^\delta$, with $\tilde{T}$ denoting the extension of T to $\mathcal{X}_s$. The bound for k_* then follows by Theorem 6.11 in [4], observing that $\mathcal{X}_u^s = \mathcal{R}((\tilde{T}^*\tilde{T})^{\frac{u-s}{2(a+s)}})$. ■

Remark 4.4 Similar as in [3], the rate can also be proven for $|\!|\!| x_{k_*}^\delta - x^\dagger |\!|\!|_0$. Together with $|\!|\!| x_{k_*}^\delta - x^\dagger |\!|\!|_{-a} = O(\delta)$ and interpolation arguments, one can derive the rates

$$|\!|\!| x_k^\delta - x^\dagger |\!|\!|_r = O(\delta^{\frac{u-r}{a+u}}), \qquad \text{for} \quad -a \leq r \leq u \tag{4.11}$$

for the intermediate spaces $\mathcal{X}_r$ with $-a \leq r \leq 0$. Observing that $\mathcal{X}_u^s = \mathcal{R}((\tilde{T}^*\tilde{T})^{\frac{u-s}{2(a+s)}})$ we see that these rates are optimal under the given source condition. If additionally the stronger condition (3.3) holds, then for $-a \leq u \leq a+2s$ the spaces $\mathcal{X}_u^s$ and $\mathcal{X}_u$ coincide with equivalent norms, and the rates (4.11) hold for $-a \leq r \leq u$ (cf. [10]).

Using the improved *a posteriori* stopping rule given in [4, Section 6], the result (4.10) holds for $0 < u \leq 2(a+s)\mu_0$, as in case of the *a priori* stopping rule.

Choosing $s = -a/2$, one has for $x^\dagger - x_0 \in \mathcal{R}((T^*T)^\mu) \cap \mathcal{X}_u^s$ with $u = 2a\mu$ the following bounds on the stopping index: $k_* = O(\delta^{\frac{2}{2\mu+1}})$ for Landweber iteration, $k_* = O(\delta^{\frac{1}{2\mu+1}})$ for $\nu-$methods (with $\nu \geq \mu + 1/2$) or Landweber iteration in Hilbert scales, and $k_* = O(\delta^{\frac{1}{2(2\mu+1)}})$ for the Hilbert scale $\nu-$methods (with $\nu \geq \mu + 1$).

For $s = 0$, Theorem 4.3 reduces to the statement of Theorem 2.1.

5 Examples and numerical tests

In this section we present several examples, where the conditions of Assumption 3.1 are satisfied and thus the results of Section 4 are applicable. We compare the performance of the proposed Hilbert-scale $\nu-$methods with standard Landweber iteration and $\nu-$methods, Landweber iteration in Hilbert scales and the method of conjugate gradients. For our numerical tests, we choose a very fine discretization by standard piecewise linear finite elements. In order to ensure that discretization effects can be neglected, we performed the test on different discretization levels, yielding almost identical results.

As a first example we consider the identification of a source term from distributed measurements:

Example 5.1 Let Ω be a bounded domain in $\mathcal{R}^n$, $n = 2, 3$ with sufficiently smooth boundary (e.g., $\partial\Omega \in \mathcal{C}^{1,1}$ or $\partial\Omega \in \mathcal{C}^{0,1}$ and Ω convex) or let Ω be a parallelepiped. Consider the operator $T : \mathcal{L}^2(\Omega) \to \mathcal{L}^2(\Omega)$ defined by $Tf = u$, with

$$Au := -\nabla \cdot (q\nabla u) + \mathbf{p} \cdot \nabla u + cu = f, \qquad u|_{\partial\Omega} = 0, \tag{5.1}$$

and given, sufficiently smooth parameters q, $\mathbf{p}$ and c. Assume that A is uniformly elliptic; then a solution to (5.1) has improved regularity $u \in \mathcal{H}^2(\Omega) \cap \mathcal{H}^1_0(\Omega)$ and satisfies $\|u\|_{\mathcal{H}^2} \sim \|f\|_{\mathcal{L}^2}$. Let $\mathcal{X}_2 = \mathcal{H}^2(\Omega) \cap \mathcal{H}^1_0(\Omega)$, with $L^2u = -\Delta u$ define the Hilbert scale $\{\mathcal{X}_s\}_{s\in\mathcal{R}}$ over $\mathcal{X} = \mathcal{L}^2(\Omega)$. Then we have $T \sim \mathcal{X}_2$, thus Assumption 3.1 holds with $a = 2$. Moreover, the stronger condition (3.3) holds.

For our numerical tests, we set $\Omega = [0, 1]^2$, $q = c = 1$ and $\mathbf{p} = 0$, and $s = -a/2 = -1$. For the $\nu-$methods we choose $\nu = 2$; note that $\nu \geq 3/2$ is necessary to apply Theorem 4.3 for $u = 2a\mu = 1/2$ in our case. We try to identify the function

$$f^\dagger = (\pi^2 + 1)\sin(\pi x) + (4\pi^2 + 1)\sin(2\pi y)$$

corresponding $u = \sin(\pi x) + \sin(2\pi y)$. As a starting value we choose $f_0 = 0$. With this setting, we have $f^\dagger \in \mathcal{R}((T^*T)^\mu)$ for all $0 \leq \mu < 1/8$, thus one would expect the iteration numbers $k_* \sim \delta^{-8/5}$ for Landweber iteration, $k_* \sim \delta^{-4/5}$ for Landweber iteration in Hilbert scales and the $\nu-$methods, and $k_* \sim \delta^{-2/5}$ for the Hilbert scale $\nu-$method. For $\Omega = [0, 1]^2$, the singular values of T behave like $\sigma_{m,n} = O((m^2 + n^2 + 1)^{-1}) = O(N^{-1})$, with $N = mn$. Thus stopping index for the conjugate gradient method can be bounded by (cf. [4, Theorem 7.14] with $\alpha = 1$)

$$k_*(cg) \leq c\delta^{-\frac{1}{(2\mu+1)(1+\alpha)}} = c\delta^{-\frac{1}{2(2\mu+1)}}, \tag{5.2}$$

which is the same bound as for the proposed Hilbert scale $\nu-$method. Finally, the error should behave like $\|f^\delta_{k_*} - f^\dagger\| \sim \delta^{1/5}$ for all methods.

$\delta/\|u\|$	$k_*(lw)$	$k_*(nu)$	$k_*(hs)$	$k_*(hsnu)$	$k_*(cg)$
0.016	86	24	11	8	6
0.008	240	42	18	10	8
0.004	725	75	29	14	13
0.002	2150	130	51	18	19
0.001	6080	219	87	24	28

Table 1. Iteration numbers for Landweber iteration (lw), the $\nu-$method (nu), Landweber iteration in Hilbert scales (hs), the proposed Hilbert scale $\nu-$method and the conjugate gradient algorithm (cg).

The numerically observed iteration numbers are $k_* = \delta^{-1.54}$ for Landweber iteration, $k_* = \delta^{-0.80}$ for the 2−method, $k_* = \delta^{-0.75}$ for Landweber iteration in Hilbert scales and $k_* = \delta^{-0.40}$ for the proposed Hilbert scale 2−method. As expected, the iteration numbers for conjugate gradients and the Hilbert scale $\nu-$method are of the same order.

Note that for $\Omega \subset \mathcal{R}^3$, the Hilbert scale method should outperform the conjugate gradient algorithm, since there we only have $\alpha = 2/3$ in (5.2) yielding $k_*(cg) \sim \delta^{-\frac{3}{5(2\mu+1)}}$ while the estimate for the semiiterative method in Hilbert scales is still $k_* \sim \delta^{-\frac{1}{2(2\mu+1)}}$. Table 2 lists the iteration error $e_{k_*} = \|f^\delta_{k_*} - f^\dagger\|$ for our numerical test:

$\delta/\|u\|$	$e_{k_*}(lw)$	$e_{k_*}(nu)$	$e_{k_*}(hs)$	$e_{k_*}(hsnu)$	$e_{k_*}(cg)$
0.016	0.328403	0.33310	0.34812	0.34210	0.32499
0.008	0.283369	0.28547	0.29510	0.28842	0.27948
0.004	0.240430	0.24163	0.25057	0.24648	0.23434
0.002	0.205203	0.20668	0.21525	0.21041	0.20361
0.001	0.175605	0.17691	0.18289	0.17889	0.17137

Table 2. Iteration error $e_{k_*} = \|f^\delta_{k_*} - f^\dagger\|$for Landweber iteration (lw), the $\nu-$method (nu), Landweber iteration in Hilbert scales (hs), the proposed Hilbert scale $\nu-$method and the conjugate gradient algorithm (cg).

The corresponding convergence rates are $e_{k_*} \sim \delta^{0.22}$ for Landweber iteration, $e_{k_*} \sim \delta^{0.23}$ for the other methods.

Originally, regularization in Hilbert scales was investigated only under the stronger condition (3.3), which is satisfied in Example 5.1. However, in the case $s \leq 0$, the condition (A2) suffices to obtain the appropriate convergence rates. In the following example, only a weaker estimate from below (3.4) holds. Note, that due to Proposition 3.2 the source condition $x^\dagger - x_0 \in \mathcal{X}_u^s$ can still be interpreted in terms of the spaces $\mathcal{X}_s$.

Consider the solution of the following Fredholm integral equation of the first kind:

Example 5.2 Let $T : \mathcal{L}^2[0,1] \to \mathcal{L}^2[0,1]$ be defined by

$$(Tx)(s) = \int_0^1 s^{1/2} k(s,t) x(t) dt, \tag{5.3}$$

with the standard Green's kernel

$$k(s,t) = \begin{cases} s(1-t), & 0 \le s < t \le 1 \\ t(1-s), & 0 \le t \le s \le 1. \end{cases}$$

Without the additional weight function $s^{1/2}$, application of the operator T would correspond to the solution of the boundary value problem $-x_{ss} = y$ with homogeneous boundary conditions (cf. Example 5.1). With

$$(T^*y)(t) = (1-t)\int_0^t s^{3/2} y(s)ds + t\int_t^1 s^{1/2}(1-s)y(s)ds$$

we get

$$\mathcal{R}(T^*) = \{w \in \mathcal{H}^2[0,1] \cap \mathcal{H}_0^1[0,1] : t^{-1/2}w''(t) \in \mathcal{L}^2[0,1]\}.$$

As Hilbert scale operator, we choose

$$L^s x := \sum_{n=1}^{\infty} (n\pi)^s \langle x, x_n \rangle x_n, \qquad x_n := \sqrt{2}\sin(n\pi\cdot), \tag{5.4}$$

and $\mathcal{L}^2 x = -x''$. This choice yields $\mathcal{R}(T^*) \subsetneq \mathcal{X}_2 := \mathcal{H}^2[0,1] \cap \mathcal{H}_0^1[0,1]$ and additionally, $\mathcal{R}(T^*) \supset \mathcal{X}_{2.5} := \{w \in \mathcal{H}^{2.5}[0,1] \cap \mathcal{H}_0^1[0,1] : \rho^{-1/2}w'' \in \mathcal{L}^2[0,1]\}$, with $\rho(t) = t(1-t)$. By Theorem 11.7 in [13], we have

$$\|w\|_{2.5}^2 \sim \|w''\|_{\mathcal{H}^{1/2}}^2 + \|\rho^{-1/2}w''\|_{\mathcal{L}^2}^2$$

and thus

$$\underline{m}\|x\|_{-2.5} \le \|Tx\| \le \overline{m}\|x\|_{-2},$$

see [3] for details.

We consider the reconstruction of the unknown function

$$x^\dagger(s) = 2t - \operatorname{sign}(2t-1) - 1,$$

and choose $s = -1$ and $x_0 = 0$. For brevity, we report only on the results obtained with Landweber iteration in Hilbert scales, the Hilbert scale $\nu-$method and the conjugate gradient algorithm:

$\delta/\|y\|$	$k_*(nu)$	$e_{k_*}(nu)$	$k_*(hsnu)$	$e_{k_*}(hsnu)$	$k_*(cg)$	$e_{k_*}(cg)$
0.016	60	0.42317	11	0.38092	5	0.38866
0.008	100	0.37634	15	0.33851	6	0.33448
0.004	178	0.32480	21	0.28897	8	0.30918
0.002	313	0.28257	28	0.25193	11	0.26304
0.001	541	0.24696	36	0.22362	14	0.23447

Table 3. Iteration numbers k_* and error $e_{k_*} = \|x_{k_*}^\delta - x^\dagger\|$ for the $2-$method (nu), the proposed Hilbert scale $2-$method (hsnu) and the conjugate gradient algorithm (cg).

The stopping indices behave like $k_* \sim \delta^{-0.8}$ for the $\nu-$method, $k_* \sim \delta^{-0.43}$ for the Hilbert scale $\nu-$method and $k_* \sim \delta^{-0.38}$ for cg. The corresponding convergence rates are $e_{k_*} \sim \delta^{0.2}$ for all examples. Again, the values are almost exactly the ones predicted by the theory ($\mu = 1/8$).

Many linear inverse problems appearing in the framework of signal and image processing, e.g., denoising or deconvolution, typically lead to Fredholm integral equations of the first kind (cf. Example 5.2) and can be treated in a similar way.

In the next example we study the problem of *Transmission Computerized Tomography* (see [16]):

Example 5.3 Let $\Omega \subset \mathcal{R}^n$, $n = 2, 3$ be a compact domain with spatially varying density f. In a simple physical model the relative intensity loss along a distance Δx is assumed to satisfy

$$\frac{\Delta I}{I} = f(x)\Delta x.$$

Denoting by $I_1(\theta, s)$ and $I_0(\theta, s)$ the intensities of the X-ray beams measured at the detector and emitter connected by the line parameterized by the distance to the origin s and the direction θ and located outside of the domain Ω, then one gets

$$(Rf)(\theta, s) := \int_{x \cdot \theta} = s f(x) dx = -\log \frac{I_1(\theta, s)}{I_0(\theta, s)}, \tag{5.5}$$

for $w \in \mathcal{R}^2$, $\|w\| = 1$ and $t > 0$. Determining the unknown density f from measurements of the intensity drop $g(\theta, s) = \frac{I_1(\theta,s)}{I_0(\theta,s)}$ hence corresponds to inversion of the *Radon transformation*. By [16, Theorem 5.1], we know that for each α there exist positive constants $c(\alpha, n)$ and $C(\alpha, n)$ such that for $f \in C_0^\infty(\Omega^n)$

$$c(\alpha, n) \|f\|_{\mathcal{H}_0^\alpha(\Omega^n)} \leq \|Rf\|_{\mathcal{H}^{\alpha+(n-1)/2}(Z)} \leq C(\alpha, n) \|f\|_{\mathcal{H}_0^\alpha(\Omega^n)},$$

with $\Omega^n \subset \mathcal{R}^n$ denoting the unit ball, and Z the cylinder $S^{n-1} \times \mathcal{R}$. This implies (3.3) for an appropriate choice of spaces; e.g., for $\mathcal{X} = \mathcal{L}^2(\Omega^n)$ and $\mathcal{Y} = \mathcal{L}^2(Z)$, we see that the Radon transformation behaves like differentiation of order one half in dimension $n = 2$, and like one times differentiation in dimension $n = 3$.

If Ω is a circle with radius r and $f(\theta, s) = f(s)$, and consequently $g(\theta, s) = g(s)$, are radially symmetric, then (5.5) can be reduced to the solution of an *Abel integral equation* of the first kind (see [16]), whose solution we investigate numerically:

Let $T : \mathcal{L}^2[0, 1] \to \mathcal{L}^2[0, 1]$ be defined by

$$(Tx)(s) := \frac{1}{\sqrt{\pi}} \int_0^s \frac{x(t)}{\sqrt{s-t}} dt, \tag{5.6}$$

with data y and "true" solution $x^\dagger = T^\dagger y$. One can show that $(T^2 x)(s) = \int_0^s x(t) dt$, thus inverting T amounts to differentiation of half order; more precisely, cf. [5],

$$\mathcal{R}(T) \subset \mathcal{H}^r[0, 1], \qquad \text{for all} \quad 0 \leq r < 1/2.$$

Let the Hilbert scale operator L be defined by

$$L^{2s} x = \sum_{n=0}^{\infty} \lambda_n^s \langle x, x_n \rangle x_n, \quad \text{with} \quad x_n(t) = \sqrt{2} \sin(\lambda_n t), \quad \lambda_n = (n + 1/2)\pi, \tag{5.7}$$

$\mathcal{X} = \mathcal{L}_2[0,1]$ and $\mathcal{X}_2 = \{x \in H^1[0,1] : x(0) = 0\}$. Then $\mathcal{R}(T) \subset \mathcal{X}_r$ for all $0 < r < 1$ and and $0 < a < 1$ and $-1/2 < s = -a/2$ is possible. Thus, the iteration can be preconditioned with L^{-a}, which corresponds to differentiation of fractional order and can be realized efficiently via (5.7) and FFT.

In the numerical test we set $s = -1/2$ (which is the limiting case of allowed choices) and try to identify the unknown density

$$x^\dagger(s) = 2t - \mathrm{sign}(2t-1) - 1.$$

The iterations are started with $x_0 = 0$. In this setting we have $x^\dagger \in X_\mu$ for all $0 \le \mu < 1/2$, thus we can expect the iteration numbers $k_* \sim \delta^{-1}$ for Landweber iteration, $k_* \sim \delta^{-1/2}$ for the $\nu-$method and Landweber iteration in Hilbert scales and $k_* \sim \delta^{-1/4}$ for the proposed Hilbert-scale $\nu-$method. The stopping index for the conjugate gradient algorithm is bounded by $k_* \sim \delta^{-1/3}$. As mentioned in the introduction, the bound for the Hilbert scale $\nu-$method is stronger than the one for the conjugate gradient algorithm if the singular values σ_n of T decay not faster than $n^{-\alpha}$ with some $0 < \alpha < 1$, which is the case here.

$\delta/\|u\|$	$k_*(lw$	$k_*(nu)$	$k_*(hs)$	$k_*(hs1)$	$k_*(hs2)$	$k_*(cg)$
0.016	37	16	9	7	6	6
0.008	75	24	12	9	8	8
0.004	146	33	15	14	10	10
0.002	300	48	21	19	12	14
0.001	643	71	31	26	15	19

Table 4. Iteration numbers k_* for the Landweber iteration, the $2-$method (nu), Landweber iteration in Hilbert scales (hs), the proposed Hilbert scale $\nu-$methods (hs1, hs2) and the conjugate gradient algorithm (cg).

The numerically realized rates for the stopping index $k_* \sim \delta^{-1.0}$ for Landweber iteration, $k_* \sim \delta^{-0.53}$ for the $2-$method, $k_* \sim \delta^{-0.44}$ for Landweber iteration in Hilbert scales and $k_* \sim \delta^{-0.4}$ for the cg-method are in good accordance with the theoretically predicted ones. The two Hilbert scales $\nu-$methods yield $k_* \sim \delta^{-0.48}$ for $\nu = 1$ and $k_* \sim \delta^{-0.3}$ for the $\nu = 2$. Note, that due to the restriction on the qualification μ_0 of the used method in Theorem 4.3, one has to choose

$$\nu \ge \frac{u-s}{2(a+s)} + \frac{1}{2} = 2,$$

in order to get optimal number of iteration and convergence rates for the Hilbert scale $\nu-$method stopped with the discrepancy principle (1.5). This explains the higher number of iterations needed for the Hilbert scale $1-$method. Finally, for all examples, the iteration error $e_{k_*} = \|x^\delta_{k_*} - x^\dagger\|$ decreases approximately like $\delta^{0.4}$ in accordance to the the predicted rate $\delta^{\frac{2\mu}{2\mu+1}}$.

In the last example, we investigate the performance of the iteration methods for an exponentially ill-posed problem: The solution of the *backwards heat equation* by Landweber iteration in Hilbert scales was already investigated in [3]. We compare the numerical results by the ones for $\nu-$methods in Hilbert scales and cg.

Example 5.4 Let $Tx = y$, $T : \mathcal{L}^2[0,1] \to \mathcal{L}^2[0,1]$ be defined by $(Tg)(x) = y(x) = u(x, \bar{t})$ for some $\bar{t} > 0$ and

$$-u_t + qu_{xx} = 0, \qquad u(0,t) = u(1,t) = 0, \qquad u(x,0) = g.$$

Let L^s be defined by (5.4). Then we have

$$\|T^* y\|_r \leq c(r)\,\|y\|_0 \qquad \text{for all} \quad r \in \mathcal{R},$$

but no estimate from below (3.4) exists.

We consider the numerical reconstruction and compare the numerically observed convergence rates and iteration numbers for the example

$$g^\dagger(x) = 2x - \operatorname{sign}(2x-1) - 1,$$

and set $g_0 = 0$. Note, that for exponentially ill-posed problems only a logarithmic convergence rate can be expected under the weak source-condition of our example.

$\delta/\|u\|$	$k_*(lw$	$k_*(nu)$	$k_*(hs)$	$k_*(hsnu)$	$k_*(cg)$
0.016	20	12	6	4	3
0.008	50	21	8	6	5
0.004	377	56	28	13	5
0.002	723	74	53	17	5
0.001	1116	88	80	21	5

Table 5. Iteration numbers k_* for the Landweber iteration, the $2-$method (nu), Landweber iteration in Hilbert scales (hs), the proposed Hilbert scale $\nu-$method (hsnu) and the conjugate gradient algorithm (cg).

The stopping indices are bounded by $k_* \sim \delta{-1.54}$ for Landweber iteration, $k_* \sim \delta^{-0.75}$ for the $\nu-$method, $k_* \sim \delta^{-1.02}$ for Landweber iteration in Hilbert scales and $k_* \sim \delta^{-0.63}$ for the Hilbert scale $\nu-$method.

According to Theorem 7.14 in [4], the stopping index for the conjugate gradient method can be bounded by $k(\delta, y^\delta) \leq c(1 + \log \frac{1}{\delta})$ for exponentially ill-posed problems, i.e. if the singular values σ_n of T decay like $O(q^n)$ with some $q < 1$. This behaviour can also be seen in the numerical test.

The numerically observed convergence rates are approximately $\|x_{k_*}^\delta - x^\dagger\| \sim \delta^{0.05}$ in all our tests.

References

[1] A. B. Bakushinskii and A. V. Goncharskii. *Iterative methods for the solution of incorrect problems.* Nauka, Moscow, 1989.

[2] H. Brakhage. On ill-posed problems and the method of conjugate gradients. In H. W. Engl and C. W. Groetsch, editors, *Inverse and Ill-posed Problems*, pages 165–175, Boston, New York, London, 1987. Academic Press.

[3] H. Egger and A. Neubauer. Preconditioning Landweber iteration in Hilbert scales. SFB-report 2004-25, Johann Radon institute for Industrial and Applied Mathematics, 2004.

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

[5] R. Gorenflo and S. Vesella. *Abel Integral Equations: Analysis and Applications.* Number 1461 in Lecture Notes in Math. Springer, Berlin, 1991.

[6] C. W. Groetsch. *Generalized Inverses of Linear Operators.* Dekker, New York, Basel, 1977.

[7] C. W. Groetsch. *The Theory of Tikhonov Regularization for Fredholm Equations of the First Kind.* Pitman, Boston, 1984.

[8] M. Hanke. Accelerated Landweber iterations for the solution of ill-posed equations. *Numer. Math.*, 60:341–373, 1991.

[9] M. Hanke, A. Neubauer, and O. Scherzer. A convergence analysis of the Landweber iteration for nonlinear ill-posed problems. *Numer. Math.*, 72:21–37, 1995.

[10] B. Kaltenbacher, A. Neubauer, and O. Scherzer. *Iterative Regularization Methods for Nonlinear Problems.* in preparation.

[11] S. G. Krein and J. I. Petunin. Scales of Banach spaces. *Russian Math. Surveys*, 21:85–160, 1966.

[12] L. Landweber. An iteration formula for Fredholm integral equations of the first kind. *Amer. J. Math.*, 73:615–624, 1951.

[13] J. L. Lions and E. Magenes. *Non-Homogeneous Boundary Value Problems and Applications: Volume I.* Springer, Berlin - Heidelberg, 1972.

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

[15] F. Natterer. Error bounds for Tikhonov regularization in Hilbert scales. *Appl. Anal.*, 18:29–37, 1984.

[16] F. Natterer. *The Mathematics of Computerized Tomography.* Teubner, Stuttgart, 1986.

[17] A. S. Nemirovskii and B. T. Polyak. Itrative methods for solving linear ill-posed problems under precise information. i. *Engrg. Cybernetics*, 22:1–11, 1984.

[18] A. Neubauer. Tikhonov regularization of nonlinear ill-posed problems in Hilbert scales. *Appl. Anal.*, 46:59–72, 1992.

[19] A. Neubauer. On Landweber iteration for nonlinear ill-posed problems in Hilbert scales. *Numer. Math.*, 85:309–328, 2000.

[20] U. Tautenhahn. Error estimates for regularization methods in Hilbert scales. *SIAM J. Numer. Anal.*, 33:2120–2130, 1996.