MATHEMATICS OF COMPUTATION
VOLUME 49, NUMBER 180
OCTOBER 1987, PAGES 523-542

An A Posteriori Parameter Choice for Ordinary and Iterated Tikhonov Regularization of Ill-Posed Problems Leading to Optimal Convergence Rates

By Helmut Gfrerer*

Abstract. We propose an a posteriori parameter choice for ordinary and iterated Tikhonov regularization that leads to optimal rates of convergence towards the best approximate solution of an ill-posed linear operator equation in the presence of noisy data. Numerical examples are given.

1. Introduction. Let X, Y be real Hilbert spaces, T: $X \to Y$ a compact linear operator, $y \in Y$. Our aim is to obtain the "best approximate solution" of

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

i.e., the unique element that has minimal norm among all minimizers of the residual $\|Tx - y\|$. If $T^\dagger$ denotes the Moore-Penrose generalized inverse (see, e.g., [17]), the best approximate solution is given by $T^\dagger y$. For nonclosed range $R(T)$ of T, the problem of determining $T^\dagger y$ is ill posed. The best approximate solution exists only for $y \in D(T^\dagger) := R(T) + R(T)^\perp$ (which we assume from now on) and depends discontinuously on the right-hand side. An important example is the (Fredholm) integral equation of the first kind

$$(Tx)(t) := \int_0^1 k(t,s)x(s)\,ds = y(t), \qquad t \in [0,1],$$

where k is a nondegenerate L^2-kernel and $X = Y = L^2[0,1]$. In the ill-posed case, the crux of the difficulty is that the data are only imprecisely known in general, that is, only some $y_\delta \in Y$ is available satisfying

$$\|y - y_\delta\| \leqslant \delta, \tag{1.2}$$

where δ is an a priori known error level. Since $T^\dagger$ is unbounded, $T^\dagger y_\delta$ is not a reasonable approximation to $T^\dagger y$, even if it exists. Because of this, one has to use "regularization methods" for approximating $T^\dagger y$. A widely used regularization method is Tikhonov regularization. For $\alpha > 0$ we denote by $x_{\alpha,\delta}$ the unique solution of

$$(\alpha I + T^*T)x = T^*y_\delta. \tag{1.3}$$

Received February 28, 1986; revised February 11, 1987.
1980 *Mathematics Subject Classification* (1985 *Revision*). Primary 65R20; Secondary 45L10.
*Parts of the work were supported by the Austrian Fonds zur Förderung der wissenschaftlichen Forschung (Project S 32/03).

It is well known (see, e.g., [3]) that if the "regularization parameter" α is chosen in dependence of δ such that $\lim_{\delta\to 0}\delta^2\alpha(\delta)^{-1} = 0$ and $\lim_{\delta\to 0}\alpha(\delta) = 0$, then $\lim_{\delta\to 0}\|x_{\alpha(\delta),\delta} - T^\dagger y\| = 0$. If the exact solution fulfills the smoothness property

$$T^\dagger y \in R\big((T^*T)^\nu\big) \tag{1.4}$$

for some $0 < \nu \leqslant 1$, then for an a priori choice of α such as

$$\alpha(\delta) = C\delta^{2/(2\nu+1)}, \quad C > 0, \tag{1.5}$$

one obtains the convergence rate

$$\left\|x_{\alpha(\delta),\delta} - T^\dagger y\right\| = O(\delta^{2\nu/(2\nu+1)}) \tag{1.6}$$

(see [19]). This convergence rate is best, $O(\delta^{2/3})$, for $\nu = 1$. A saturation result of Groetsch [11] says that a higher rate of convergence cannot be expected under higher smoothness assumptions and other choices of $\alpha(\delta)$. However, a higher convergence rate can be obtained by "iterated Tikhonov regularization" (see [14]), which is defined by the formulas

$$x^0_{\alpha,\delta} := 0; \qquad (\alpha I + T^*T)x^j_{\alpha,\delta} = T^*y_\delta + \alpha x^{j-1}_{\alpha,\delta}, \qquad j = 1,\dots,n. \tag{1.7}$$

If the smoothness condition (1.4) holds for some $0 < \nu \leqslant n$, then a parameter choice according to (1.5) yields a convergence rate

$$\left\|x^n_{\alpha(\delta),\delta} - T^\dagger y\right\| = O(\delta^{2\nu/(2\nu+1)}) \tag{1.8}$$

which is best, $O(\delta^{2n/(2n+1)})$, for $\nu = n$. Unfortunately, one cannot determine $\alpha(\delta)$ by (1.5) in practice, since the number ν depends on the unknown solution $T^\dagger y$. Therefore, many authors suggest a posteriori methods to compute a reasonable value of α using the input data y_δ and the error level δ. A favorite choice of α is the so-called "discrepancy principle" due to Morozov [16], where $\alpha = \alpha(\delta)$ is computed as the unique solution of

$$\|Tx_{\alpha,\delta} - y_\delta\|^2 = \delta^2. \tag{1.9}$$

Arcangeli [1] proposes $\alpha = \alpha(\delta)$ as solution of

$$\|Tx_{\alpha,\delta} - y_\delta\|^2 = \delta^2\alpha^{-1}, \tag{1.10}$$

while Engl [4] (for a similar method; see Schock [20]) suggests choosing $\alpha = \alpha(\delta)$ as the unique root of

$$\|T^*Tx_{\alpha,\delta} - T^*y_\delta\|^2 = \delta^p\alpha^{-q} \tag{1.11}$$

with suitable constants p, q. Engl [5] applied his method also to iterated Tikhonov regularization.

All these methods do not yield the convergence rates given by (1.6). For Morozov's, resp. Arcangeli's approach, this is shown in [10], resp. [13]. Engl has to choose the parameters p and q in (1.11) in dependence of the unknown quantity ν to obtain the rates (1.6).

The aim of this paper is to give an a posteriori method for choosing the regularization parameter for iterated Tikhonov regularization, where no information about ν is used and the rates (1.8) are achieved, and even improved upon, for $\nu < n$. The difference is the replacement of the capital-O condition by the little-o condition.

The basic idea of our method is rather simple. Obviously, the best possible parameter choice would be such that the squared error $\|x^n_{\alpha,\delta} - T^\dagger y\|^2$ is minimized. Of course, this criterion is not applicable, but we will find a minimizer of some upper bound of the squared error.

At the end of Section 2 we investigate convergence rates for Morozov's discrepancy principle (1.9). It is well known (see [12]) that a certain upper bound of the squared error is minimized precisely when the parameter is chosen according to (1.9). Using the same technique of proof as for our method, we will show that Morozov's method yields also the convergence rates (1.6), but only for $\nu \leqslant 1/2$. In Section 3 we adapt the theory developed in Section 2 to make it applicable to practical computations. For this purpose, we consider approximations to the best approximate solution which lie in a finite-dimensional subspace V_m of X. More precisely, for each $\alpha > 0$ and $n \in \mathbf{N}$, we define $x^n_{\alpha,\delta,m}$ iteratively by the formulas

$$(1.12) \quad x^0_{\alpha,\delta,m} := 0; \qquad \left(\alpha I + T_m^* T_m\right) x^j_{\alpha,\delta,m} = T_m^* y_\delta + \alpha x^{j-1}_{\alpha,\delta,m}, \qquad j = 1, \dots, n,$$

where $T_m := TP_m$ and P_m is the orthogonal projector of X onto V_m. For $n = 1$, this is equivalent to the approaches of Groetsch [12], Engl and Neubauer [6] and closely related to Marti's method [15].

Now the regularization parameter has to be chosen appropriately in dependence of the noise level δ and the subspace V_m. So Groetsch [12] applied the discrepancy principle to this finite-dimensional setting, whereas Engl and Neubauer [6] modified (1.11) to obtain finite-dimensional approximations. In view of known results, these methods seem to have the same disadvantages mentioned above for the infinite-dimensional case.

We give in Section 3 an a posteriori parameter choice $\alpha = \alpha(\delta, V_m, n)$ such that for $T^\dagger y \in R((T^*T)^\nu)$ we have

$$\left\|x^n_{\alpha,\delta,m} - T^\dagger y\right\| = \begin{cases} o\left(\delta^{2\nu/(2\nu+1)}\right) + o\left(\gamma_m^{2\nu}\right) & \text{if } 0 < \nu < 1/2, \\ o\left(\delta^{2\nu/(2\nu+1)}\right) + O(\gamma_m) & \text{if } 1/2 \leqslant \nu < n, \\ O\left(\delta^{2n/(2n+1)}\right) + O(\gamma_m) & \text{if } \nu \geqslant n, \end{cases}$$

where $\gamma_m = \|T(I - P_m)\|$ is a measure of how well T_m approximates T. Again, this method requires no information about ν and is numerically feasible in the sense that it depends on finitely many numerical parameters.

In Section 4 numerical examples are given which show that theory and practice agree quite well.

2. Optimal Parameter Choice for Iterated Tikhonov Regularization with Inexact Data. From now on we assume that $y \in D(T^\dagger)$, $T^\dagger y \neq 0$, and we wish to determine $x = T^\dagger y$, having at our disposal only an approximation y_δ satisfying

$$(2.1) \qquad \|y - y_\delta\|^2 \leqslant \delta^2 < \|Qy_\delta\|^2 / C$$

for some $C \geqslant 1$, where Q denotes the orthogonal projector onto $\overline{R(T)}$. We believe that condition (2.1) is not a severe restriction. It may be interpreted by saying that the relative error in the input data, $\delta/\|Qy_\delta\|$, is small. For $n \in \mathbf{N}$, $z \in Y$, $\alpha > 0$ we define

$$(2.2) \qquad f_n(\alpha, z) := \alpha^{2n+1}\left\langle (\alpha I + TT^*)^{-(2n+1)} Qz, Qz \right\rangle.$$

We further define the error functions

$$(2.3)\qquad \varphi_n(\alpha) := \left\| x_\alpha^n - T^\dagger y \right\|^2, \qquad \varphi_{n,\delta}(\alpha) = \left\| x_{\alpha,\delta}^n - T^\dagger y \right\|^2,$$

where x_α^n is given by (1.7) with y instead of y_δ. The following lemma states some properties of these functions. Proofs of this and subsequent results are given in appendices in the Supplements section of this issue.

LEMMA 2.1. *For each $z \in Y$ with $Qz \neq 0$ and each $n \in \mathbf{N}$, the function $\alpha \to f_n(\alpha, z)$ is continuous, strictly increasing on $(0, \infty)$ and $\lim_{\alpha\to 0} f_n(\alpha, z) = 0$, $\lim_{\alpha\to\infty} f_n(\alpha, z) = \|Qz\|^2$. For each $n \in \mathbf{N}$, the function $\alpha \to \varphi_n(\alpha)$ is continuously differentiable and strictly increasing on $(0, \infty)$. Furthermore, there holds*

$$\alpha^2 \varphi_n'(\alpha) = 2n f_n(\alpha, y), \quad \lim_{\alpha\to 0} \varphi_n(\alpha) = 0, \quad \lim_{\alpha\to\infty} \varphi_n(\alpha) = \left\| T^\dagger y \right\|^2.$$

Proof. Cf. Appendix 2. □

Application of the first part of Lemma 2.1 and the Intermediate Value Theorem yields the following corollary.

COROLLARY 2.2. *Suppose that y, $y_\delta \in Y$ and $\delta > 0$ satisfy (2.1) for some $C \geqslant 1$. Then for each $n \in \mathbf{N}$ there is a unique $\alpha > 0$ such that*

$$(2.4)\qquad f_n(\alpha, y_\delta) = C\delta^2.$$

From now on, we denote the unique α determined by (2.4) by $\alpha(\delta)$ (although it depends also on y_δ, C, and n). An expression involving $\alpha(\delta)$ will be understood in the sense that for fixed $C \geqslant 1$ and $n \in \mathbf{N}$ it holds for all y_δ satisfying (2.1) and the corresponding α determined by (2.4).

The next theorem gives some motivation for our proposed parameter choice.

THEOREM 2.3. *Let $n \in \mathbf{N}$ be fixed and suppose that y, $y_\delta \in Y$ and $\delta > 0$ satisfy (2.1) for some $C \geqslant 1$. Then the function $\varphi_{n,\delta}(\alpha)$ is strictly increasing for $\alpha > \bar{\alpha}$, where $\bar{\alpha} > 0$ is the unique solution of $f_n(\alpha, y_\delta) = \delta^2$.*

Proof. Using (2.3) and (1.7), we have

$$\varphi_{n,\delta}(\alpha) = \left\| \sum_{k=1}^{n} \alpha^{k-1}(\alpha I + T^*T)^{-k} T^* y_\delta - T^\dagger y \right\|^2.$$

It is easy to see that $\varphi_{n,\delta}(\alpha)$ is continuously differentiable for $\alpha > 0$ and

$$\begin{aligned} \varphi_{n,\delta}'(\alpha) &= 2\Big\langle \sum_{k=1}^{n} \alpha^{k-1}(\alpha I + T^*T)^{-k} T^* y_\delta - T^\dagger y, \\ &\qquad \sum_{k=1}^{n} \Big[(k-1)\alpha^{k-2}(\alpha I + T^*T)^{-k} T^* y_\delta \\ &\qquad\qquad - k\alpha^{k-1}(\alpha I + T^*T)^{-k-1} T^* y_\delta\Big]\Big\rangle \qquad (2.5) \\ &= -2n\Big\langle \sum_{k=1}^{n} \alpha^{k-1}(\alpha I + TT^*)^{-k} TT^* y_\delta - Qy,\ \alpha^{n-1}(\alpha I + TT^*)^{-(n+1)} Q y_\delta \Big\rangle \\ &= 2n\Big\langle \alpha^n(\alpha I + TT^*)^{-n} Q y_\delta - Q y_\delta + Qy,\ \alpha^{n-1}(\alpha I + TT^*)^{-(n+1)} Q y_\delta \Big\rangle. \end{aligned}$$

To prove our theorem, it suffices to show that $\varphi'_{n,\delta}(\alpha) > 0$ for $\alpha > \bar{\alpha}$. But $\alpha > \bar{\alpha}$ implies by Lemma 2.1 that $f_n(\alpha, y_\delta) > f_n(\bar{\alpha}, y_\delta) = \delta^2$ and hence

$$f_n(\alpha, y_\delta) > \delta f_n(\alpha, y_\delta)^{1/2} \geqslant \|Q(y_\delta - y)\| f_n(\alpha, y_\delta)^{1/2}. \tag{2.6}$$

Since $\|\alpha^{1/2}(\alpha I + TT^*)^{-1/2}\| \leqslant 1$, we obtain

$$f_n(\alpha, y_\delta)^{1/2} = \left\|\alpha^{n+1/2}(\alpha I + TT^*)^{-(n+1/2)} Q y_\delta\right\| \geqslant \left\|\alpha^{n+1}(\alpha I + TT^*)^{-(n+1)} Q y_\delta\right\|.$$

This, together with (2.6), yields for $\alpha > \bar{\alpha}$,

$$\begin{aligned} f_n(\alpha, y_\delta) &> \|Qy_\delta - Qy\| \cdot \left\|\alpha^{n+1}(\alpha I + TT^*)^{-(n+1)} Q y_\delta\right\| \\ &\geqslant \left\langle Qy_\delta - Qy, \alpha^{n+1}(\alpha I + TT^*)^{-(n+1)} Q y_\delta\right\rangle. \end{aligned}$$

By (2.5) this is equivalent to $(\alpha^2/2n)\varphi'_{n,\delta}(\alpha) > 0$, and hence our theorem is proved. □

Remark 2.4. Theorem 2.3 and the first part of Lemma 2.1 show that a choice $C = 1$ in (2.4) yields the best result among all possible choices of $C \geqslant 1$. However, for technical reasons, we also have to consider the case $C > 1$.

Our convergence analysis is mainly based on the following three lemmas, whose proofs are given in Appendix 2.

LEMMA 2.5. *Let $C > 1$, $C_1 := (C^{1/2} - 1)^2$, $C_2 := (1 + C^{1/2})^2$. Then for each $\delta > 0$, y, $y_\delta \in Y$ satisfying* (2.1), *and for each $n \in \mathbf{N}$, we have*

$$C_1\delta^2 \leqslant f_n(\alpha(\delta), y) \leqslant C_2\delta^2.$$

LEMMA 2.6. *Let $\gamma > 0$. Then for all $\alpha > 0$ and all $n \in \mathbf{N}$ there holds*

$$\varphi_{n,\delta}(\alpha) \leqslant 2/\min\{\gamma, 1\}\big(2\gamma n\delta^2/\alpha + \varphi_n(\alpha)\big).$$

LEMMA 2.7. *Let $\gamma > 0$, $\delta > 0$, $n \in \mathbf{N}$. Then $\bar{\alpha}$ is a minimizer for the one-dimensional optimization problem*

$$\begin{aligned} &\textit{minimize } 2n\gamma\delta^2/\alpha + \varphi_n(\alpha) \\ &\textit{subject to } \alpha > 0 \end{aligned} \tag{2.7}$$

if and only if $f_n(\bar{\alpha}, y) = \gamma\delta^2$ holds.

We are now in a position to derive rates of convergence for our parameter choice.

THEOREM 2.8. *Let $C \geqslant 1$, $n \in \mathbf{N}$ be fixed. For each $\delta > 0$ and $y_\delta \in Y$ satisfying* (2.1), *let $x^n_{\alpha(\delta),\delta}$ be the result of iterated Tikhonov regularization of order n as described by* (1.7), *where $\alpha(\delta)$ is the unique solution of* (2.4). *Then $\lim_{\delta \to 0} x^n_{\alpha(\delta),\delta} = T^\dagger y$. If, further, $T^\dagger y$ is an element of $R((T^*T)^\nu)$ with $\nu > 0$, then*

$$\left\|x^n_{\alpha(\delta),\delta} - T^\dagger y\right\| = o\left(\delta^{2\nu/(2\nu+1)}\right) \quad \textit{for } \nu < n$$

and

$$\left\|x^n_{\alpha(\delta),\delta} - T^\dagger y\right\| = O\left(\delta^{2n/(2n+1)}\right) \quad \textit{for } \nu \geqslant n.$$

Proof. First suppose $C > 1$; let C_1, C_2 be as in Lemma 2.5 and set $\gamma := f_n(\alpha(\delta), y)/\delta^2$. Thus $C_1 \leqslant \gamma \leqslant C_2$. By Lemma 2.7, $\alpha(\delta)$ is a minimizer for $2n\gamma\delta^2/\alpha + \varphi_n(\alpha)$. Thus we obtain, by Lemma 2.6,

$$\begin{aligned} \varphi_{n,\delta}(\alpha(\delta)) &\leqslant 2/\min\{\gamma, 1\}\big(2n\gamma\delta^2/\alpha(\delta) + \varphi_n(\alpha(\delta))\big) \\ &= 2/\min\{\gamma, 1\}\min\big\{2n\gamma\delta^2/\alpha + \varphi_n(\alpha) : \alpha > 0\big\} \\ &\leqslant 2/\min\{C_1, 1\}\inf\big\{2nC_2\delta^2/\alpha + \varphi_n(\alpha) : \alpha > 0\big\}. \end{aligned} \tag{2.8}$$

Using Lemma 2.1 and the Intermediate Value Theorem, it can be easily shown that for each $\delta > 0$ the equation

$$\varphi_n(\beta) = 2n\delta^2/\beta \tag{2.9}$$

has a unique solution, which we denote by $\beta_n(\delta)$. Since $\beta_n(\delta)\varphi_n(\beta_n(\delta)) = 2n\delta^2$, we have, together with Lemma 2.1, $\lim_{\delta\to 0}\beta_n(\delta) = 0$ and also $\lim_{\delta\to 0} 2n\delta^2/\beta_n(\delta) = \lim_{\delta\to 0}\varphi_n(\beta_n(\delta)) = 0$. Because of (2.8)

$$\varphi_{n,\delta}(\alpha) \leqslant 2/\min\{C_1, 1\}\left(2nC_2\delta^2/\beta_n(\delta) + \varphi_n(\beta_n(\delta))\right), \tag{2.10}$$

and hence $\lim_{\delta\to 0} x^n_{\alpha(\delta),\delta} = T^\dagger y$.

Now assume that $T^\dagger y \in R((T^*T)^\nu)$ with $\nu > 0$ and set $\mu := \min\{\nu, n\}$. Then we get, from (2.9),

$$\beta_n(\delta)^{1+2\mu} = 2n\delta^2\beta_n(\delta)^{2\mu}/\varphi_n(\beta_n(\delta))$$

and hence

$$\beta_n(\delta) = (2n\delta^2)^{1/(1+2\mu)}\left[\beta_n(\delta)^{2\mu}/\varphi_n(\beta_n(\delta))\right]^{1/(1+2\mu)}.$$

Thus we have

$$2n\delta^2/\beta_n(\delta) = (2n\delta^2)^{2\mu/(1+2\mu)}\left[\varphi_n(\beta_n(\delta))/\beta_n(\delta)^{2\mu}\right]^{1/(1+2\mu)}.$$

We obtain from [18] that $\varphi_n(\beta) = o(\beta^{2\mu})$ for $\mu < n$, resp. $\varphi_n(\beta) = O(\beta^{2n})$ for $\mu = n$. This, together with $\lim_{\delta\to 0}\beta_n(\delta) = 0$, shows that

$$2n\delta^2/\beta_n(\delta) = \varphi_n(\beta_n(\delta)) = o(\delta^{4\mu/(1+2\mu)}) \quad \text{for } \nu = \mu < n,$$

resp.

$$2n\delta^2/\beta_n(\delta) = \varphi_n(\beta_n(\delta)) = O(\delta^{4\mu/(1+2\mu)}) \quad \text{for } \nu \geqslant n = \mu.$$

Because of (2.10) this implies our assertion for $C > 1$. For $C = 1$ the result follows from Remark 2.4. □

Theorem 2.8 says that the convergence rate can be arbitrarily close to the desirable rate $O(\delta)$ if the data are sufficiently smooth and n is chosen sufficiently large. However, our upper bound of the squared error, given by Lemma 2.6, involves the factor n, and so it might be problematic to choose n too large. In particular, for given $y_\delta \in Y$ and $\delta > 0$, the approximation error $\|x^n_{\alpha(\delta),\delta} - T^\dagger y\|$ might become arbitrarily large if n tends to ∞. However, our next theorem shows that this is not the case.

THEOREM 2.9. *For $C > 1$ let C_1, C_2 be as in Lemma* 2.5. *Then for each $y_\delta \in Y$, $\delta > 0$ satisfying* (2.1), *and for each $n \in \mathbf{N}$, there holds*

$$\varphi_{n,\delta}(\alpha(\delta)) \leqslant 2(C_2 + 1)/\min\{C_1, 1\}\,\varphi_n(\beta_n(\delta)),$$

where $\beta_n(\delta)$ is given by (2.9). *Further, if $m < n$, then $\varphi_n(\beta_n(\delta)) \leqslant \varphi_m(\beta_m(\delta))$ for all $\delta > 0$.*

Proof. Cf. Appendix 2. □

We now show that Morozov's discrepancy principle (1.9), under the smoothness assumption $T^\dagger y \in R((T^*T)^\nu)$, yields the convergence rates $o(\delta^{2\nu/(2\nu+1)})$ for $\nu < 1/2$, resp. $O(\delta^{1/2})$ for $\nu \geqslant 1/2$. More precisely, suppose that

$$y \in R(T), \tag{2.11}$$

and that our approximate right-hand side y_δ satisfies

$$\|y - y_\delta\| < \delta < \|y_\delta\|. \tag{2.12}$$

Then it is well known (see, e.g., [12, Theorem 3.3.1]) that (1.9) has a unique positive solution, which we denote again by $\alpha(\delta)$. Note that

$$\begin{aligned} Tx_{\alpha,\delta} - y_\delta &= TT^*(\alpha I + TT^*)^{-1}Qy_\delta - Qy_\delta - (I - Q)y_\delta \\ &= -\alpha(\alpha I + TT^*)^{-1}Qy_\delta - (I - Q)y_\delta, \end{aligned} \tag{2.13}$$

and hence $T^*(y_\delta - Tx_{\alpha,\delta}) = \alpha x_{\alpha,\delta}$. Thus we obtain

$$\begin{aligned} \|x_{\alpha,\delta} - T^\dagger y\|^2 &= \|x_{\alpha,\delta}\|^2 - 2/\alpha\langle T^*(y_\delta - Tx_{\alpha,\delta}), T^\dagger y\rangle + \|T^\dagger y\|^2 \\ &= \|x_{\alpha,\delta}\|^2 - 2/\alpha\langle y_\delta - Tx_{\alpha,\delta}, y\rangle + \|T^\dagger y\|^2 \\ &= \|x_{\alpha,\delta}\|^2 - 2/\alpha\langle y_\delta - Tx_{\alpha,\delta}, y_\delta\rangle + 2/\alpha\langle y_\delta - Tx_{\alpha,\delta}, y_\delta - y\rangle + \|T^\dagger y\|^2. \end{aligned}$$

Since

$$-\delta\|y_\delta - Tx_{\alpha,\delta}\| \leqslant \langle y_\delta - Tx_{\alpha,\delta}, y_\delta - y\rangle \leqslant \delta\|y_\delta - Tx_{\alpha,\delta}\|,$$

we obtain

$$E(\alpha, y_\delta) - 4\delta/\alpha\|y_\delta - Tx_{\alpha,\delta}\| \leqslant \|x_{\alpha,\delta} - T^\dagger y\|^2 \leqslant E(\alpha, y_\delta),$$

where

$$E(\alpha, y_\delta) := \|x_{\alpha,\delta}\|^2 - 2/\alpha\langle y_\delta - Tx_{\alpha,\delta}, y_\delta\rangle + 2\delta/\alpha\|y_\delta - Tx_{\alpha,\delta}\| + \|T^\dagger y\|^2.$$

Thus,

$$\|x_{\alpha,\delta} - T^\dagger y\|^2 \leqslant E(\alpha, y_\delta) \leqslant \|x_{\alpha,\delta} - T^\dagger y\|^2 + 4\delta/\alpha\|y_\delta - Tx_{\alpha,\delta}\|.$$

Because of $\|(I - Q)y_\delta\| = \|(I - Q)(Qy - y_\delta)\| = \|(I - Q)(y - y_\delta)\| < \delta$ and $\|(\alpha I + TT^*)^{-1}(Qy_\delta - y)\| \leqslant \delta/\alpha$, we obtain, together with (2.13),

$$\begin{aligned} \|x_{\alpha,\delta} - T^\dagger y\|^2 \leqslant E(\alpha, y_\delta) &\leqslant \|x_{\alpha,\delta} - T^\dagger y\|^2 \\ &\quad + 4\delta\|(\alpha I + TT^*)^{-1}y\| + 8\delta^2/\alpha. \end{aligned} \tag{2.14}$$

It is well known ([12, Theorem 3.3.2]) that $E(\alpha, y_\delta)$ is a minimum if and only if $\alpha = \alpha(\delta)$. Hence, we may use a similar proof to that of Theorem 2.8 to establish convergence rates for Morozov's discrepancy principle.

THEOREM 2.10. *For each* $\delta > 0$ *and* $y_\delta \in Y$ *satisfying* (2.12), *let* $\alpha(\delta)$ *be the unique solution of* (1.9). *If* $T^\dagger y \in R((T^*T)^\nu)$ *for some* $\nu > 0$, *then*

$$\|x_{\alpha(\delta),\delta} - T^\dagger y\| = o(\delta^{2\nu/(2\nu+1)}) \quad \textit{for } \nu < 1/2$$

and

$$\|x_{\alpha(\delta),\delta} - T^\dagger y\| = O(\delta^{1/2}) \quad \textit{for } \nu \geqslant 1/2.$$

Proof. Cf. Appendix 2. □

Remark 2.11. The discrepancy principle (1.9) so far described is only applicable if $y \in R(T)$. However, if $\alpha(\delta)$ is computed as the unique solution of

$$\|Tx_{\alpha,\delta} - Qy_\delta\|^2 = \delta^2,$$

then Theorem 2.10 also holds for $y \in R(T) + R(T)^\perp = D(T^\dagger)$.

3. Finite-Dimensional Approximations of Best Approximate Solutions. Let $V_1 \subseteq V_2 \subseteq V_3 \subseteq \cdots$ be a sequence of finite-dimensional subspaces of X with $\overline{\bigcup_{m \in \mathbf{N}} V_m} = X$. For each $m \in \mathbf{N}$ let

$$(3.1) \qquad T_m := TP_m, \quad \gamma_m := \|T(I - P_m)\|, \quad b_m := \|Q_m T(I - P_m)\|,$$

where P_m is the orthogonal projector of X onto V_m and Q_m is the orthogonal projector of Y onto $R(T_m)$, the range of T_m. For given $m, n \in \mathbf{N}$, $\alpha > 0$ and $y_\delta \in Y$ we consider approximations $x^n_{\alpha,\delta,m}$ given by (1.12), which lie in $R(T_m^*) \subset V_m$. In this finite-dimensional setting we assume that the available data y_δ satisfies, with an a priori given noise level $\delta \geqslant 0$,

$$(3.2) \qquad \|y - y_\delta\|^2 \leqslant \delta^2 < \|Q_m y_\delta\|^2 / C,$$

where $C > 1$. Note that for this finite-dimensional approach also the case $\delta = 0$ (i.e., the data are exactly known) is of interest, because it is not always possible to guarantee convergence of $T_m^\dagger y \to T^\dagger y$ if $m \to \infty$ (see Seidman [21]).

Our first result shows that $T_m^\dagger y$ is the best possible approximation of $T^\dagger y$ by elements of V_m, if $b_m = 0$.

THEOREM 3.1. *If $b_m = 0$, then $T_m^\dagger y = P_m T^\dagger y$.*

Proof. By (3.1) we have $Q_m T = Q_m T P_m = T_m$. Hence $T_m^\dagger y = T_m^\dagger Q_m Q y = T_m^\dagger Q_m T T^\dagger y = T_m^\dagger T_m T^\dagger y$. $T_m^\dagger T_m$ is the orthogonal projector onto $N(T_m)^\perp \subset V_m$, P_m is the orthogonal projector onto V_m and $T^\dagger y \in N(T)^\perp \subset (N(T) \cap V_m)^\perp = (N(T_m) \cap V_m)^\perp$. Therefore, we have $T_m^\dagger T_m T^\dagger y = P_m T^\dagger y$. $\square$

For convergence rates in the situation of Theorem 3.1, see Theorem 3.6. From now on we will assume that the number b_m and the noise level δ are not both zero. Then, for $n, m \in \mathbf{N}$ and $z \in Y$, we define

$$(3.3) \qquad f_n^m(\alpha, z) := \alpha^{2n+1}\left\langle \left(\alpha I + T_m T_m^*\right)^{-(2n+1)} Q_m z, Q_m z\right\rangle \quad \text{for } \alpha > 0.$$

LEMMA 3.2. *Let $K > 0$ and suppose that y, y_0 satisfy* (3.2) *for some $C > 1$. Further, assume that b_m and δ are not both zero. Then for each $n \in \mathbf{N}$ the equation*

$$(3.4) \qquad \left(1 - K(2n-1)b_m^2/\alpha\right) f_n^m(\alpha, y_\delta) = C\delta^2$$

has a unique solution $\alpha > 0$.

Proof. Cf. Appendix 2. $\square$

From now on, we denote the unique solution of (3.4) by $\alpha_m(\delta)$. We will show that a parameter choice $\alpha = \alpha_m(\delta)$ yields the convergence rates mentioned in Section 1. For our convergence analysis we may assume without loss of generality that $\gamma_m > 0$. For, if $\gamma_m = 0$, then $T = T_m$, $b_m = 0$, and so we have the same situation as discussed in Section 2.

Let $x^n_{\alpha,m}$ be the result of (1.12) with y_δ replaced by y. Then, analogously to Lemma 2.6, one can show the following lemma.

LEMMA 3.3. *Let $\eta > 0$. Then for all $\alpha > 0$ and all $n \in \mathbf{N}$ there holds*

$$\left\|x^n_{\alpha,\delta,m} - T^\dagger y\right\|^2 \leqslant 2/\min\{\eta,1\}\left(2n\eta\delta^2/\alpha + \left\|x^n_{\alpha,m} - T^\dagger y\right\|^2\right).$$

To simplify notations, we define for $j \geqslant 1$ and $\alpha > 0$

$$z^j_{\alpha,m} := \alpha^{j-1}\left(\alpha I + T_m^* T_m\right)^{-j} T_m^* y, \qquad z^j_\alpha := \alpha^{j-1}\left(\alpha I + T^*T\right)^{-j} T^* y.$$

LEMMA 3.4. *For all $\alpha > 0$ and all $n \in \mathbf{N}$ there holds*

$$\left\|x^n_{\alpha,m} - x^n_\alpha\right\| \leqslant \sum_{j=1}^{n} (n-j+1)\left(1 + b_m\alpha^{-1/2}/2\right)\left\|(I - P_m) z^j_\alpha\right\|.$$

Proof. Cf. Appendix 2. □

LEMMA 3.5. *Let $n \in \mathbf{N}$ be fixed. Then there exists a function g with $\lim_{t\to 0} g(t) = 0$ such that for all $\alpha \geqslant \gamma_m^2$,*

$$\left\|x^n_{\alpha,m} - x^n_\alpha\right\| \leqslant g(\gamma_m).$$

If further $T^\dagger y \in R((T^*T)^\nu)$ for some $\nu > 0$, then

$$g(\gamma_m) = \begin{cases} o\left(\gamma_m^{2\nu}\right) & \text{for } \nu < 1/2, \\ O(\gamma_m) & \text{for } \nu \geqslant 1/2. \end{cases}$$

Proof. Cf. Appendix 2. □

The next result gives a convergence rate in terms of γ_m for the best possible approximation of $T^\dagger y$ by elements of V_m.

THEOREM 3.6. *If $T^\dagger y \in R((T^*T)^\nu)$ for some $\nu > 0$, then*

$$\left\|(I - P_m) T^\dagger y\right\| = \begin{cases} o\left(\gamma_m^{2\nu}\right) & \text{for } \nu < 1/2, \\ O(\gamma_m) & \text{for } \nu \geqslant 1/2. \end{cases}$$

Proof. Since

$$\begin{aligned} \left\|(I - P_m) T^\dagger y\right\| &\leqslant \left\|x^1_{\alpha,m} - T^\dagger y\right\| \\ &\leqslant \left\|x^1_{\alpha,m} - x^1_\alpha\right\| + \left\|x^1_\alpha - T^\dagger y\right\| \quad \text{for all } \alpha > 0 \end{aligned} \tag{3.5}$$

and

$$\left\|x^1_\alpha - T^\dagger y\right\| = \begin{cases} o(\alpha^\nu) & \text{for } \nu < 1, \\ O(\alpha) & \text{for } \nu \geqslant 1 \end{cases}$$

(see [19]), the result follows from Lemma 3.5, if we choose $\alpha = \gamma_m^2$ in (3.5). □

In order to present the main convergence theorem, we need the following lemmas.

LEMMA 3.7. *For each $K > 0$ and each $\alpha > 0$, $n \in \mathbf{N}$, there holds*

$$\begin{aligned} \left\|x^n_{\alpha,m} - T^\dagger y\right\|^2 &\leqslant \left\|x^n_{\alpha,m} - T_m^\dagger y\right\|^2 + K b_m^2 \sum_{j=1}^{2n-1} j\rho_{j,m}(\alpha) + (1/K)\left\|(I - P_m) T^\dagger y\right\|^2 \\ &\quad + \left\|T_m^\dagger y - T^\dagger y\right\|^2 - 2\left\langle T_m^\dagger y, T_m^\dagger y - T^\dagger y\right\rangle \\ &\leqslant \left\|x^n_{\alpha,m} - T^\dagger y\right\|^2 + 2K b_m^2 \sum_{j=1}^{2n-1} j\rho_{j,m}(\alpha) + (2/K)\left\|(I - P_m) T^\dagger y\right\|^2, \end{aligned}$$

where $\rho_{j,m}(\alpha) := \alpha^{j-1}\langle(\alpha I + T_m T_m^*)^{-(j+1)} Q_m y, Q_m y\rangle$.

Proof. Cf. Appendix 2. □

LEMMA 3.8. *Let $n \in \mathbf{N}$ be fixed. Then there exists a function h with* $\lim_{t \to 0} h(t) = 0$ *such that for all* $\alpha \geqslant \gamma_m^2$,

$$b_m^2 \sum_{j=1}^{2n-1} j\rho_{j,m}(\alpha) \leqslant h(\gamma_m).$$

If further $T^\dagger y \in R((T^*T)^\nu)$ *for some* $\nu > 0$, *then*

$$h(\gamma_m) = \begin{cases} o(\gamma_m^{4\nu}) & \text{for } \nu < 1/2, \\ O(\gamma_m^2) & \text{for } \nu \geqslant 1/2. \end{cases}$$

Proof. Cf. Appendix 2. □

The proof of the following lemma is analogous to that of Lemma 2.5 and is omitted.

LEMMA 3.9. *Let* $K > 0$, $C > 1$, $C_1 := (C^{1/2} - 1)^2$, $C_2 := (C^{1/2} + 1)^2$. *Then for each* $\delta \geqslant 0$, $y, y_\delta \in Y$ *satisfying* (3.2), *and for each* $n \in \mathbf{N}$, *we have*

$$C_1\delta^2 \leqslant (1 - K(2n-1)b_m^2/\alpha_m(\delta)) f_n^m(\alpha_m(\delta), y) \leqslant C_2\delta^2,$$

where $\alpha_m(\delta)$ *is given by* (3.4).

Using Lemma 3.9 and the next lemma, we will see that our parameter choice (3.4) gives the minimum of an upper bound for the squared error $\|x_{\alpha,\delta,m}^n - T^\dagger y\|^2$.

LEMMA 3.10. *Let* $\eta > 0$, $\delta \geqslant 0$, $K > 0$, $n \in \mathbf{N}$, *and assume that* b_m *and* δ *are not both zero. Then* $\bar\alpha$ *is a minimizer for the one-dimensional optimization problem*

$$\text{minimize } 2n\eta\delta^2/\alpha + \left\|x_{\alpha,m}^n - T_m^\dagger y\right\|^2 + Kb_m^2 \sum_{j=1}^{2n-1} j\rho_{j,m}(\alpha)$$
$$+ (1/K)\left\|(I - P_m)T^\dagger y\right\|^2 - 2\left\langle T_m^\dagger y, T_m^\dagger y - T^\dagger y\right\rangle + \left\|T_m^\dagger y - T^\dagger y\right\|^2$$

subject to $\alpha > 0$

if and only if

$$(1 - K(2n-1)b_m^2/\bar\alpha) f_n^m(\bar\alpha, y) = \eta\delta^2.$$

Proof. Cf. Appendix 2. □

THEOREM 3.11. *Let* $C > 1$, $n \in \mathbf{N}$, $K > 0$ *be fixed. For each* $\delta \geqslant 0$, $y_\delta \in Y$ *satisfying* (3.2), *and for each* $V_m \subset X$ *such that* b_m *and* δ *are not both zero, let* $x_{\alpha_m(\delta),\delta,m}^n$ *be given by* (1.12), *where* $\alpha_m(\delta)$ *is the unique positive solution of* (3.4). *Then* $\lim_{\delta \to 0, m \to \infty} x_{\alpha_m(\delta),\delta,m}^n = T^\dagger y$. *If further* $T^\dagger y$ *is an element of* $R((T^*T)^\nu)$ *for some* $\nu > 0$, *then*

$$\left\|x_{\alpha_m(\delta),\delta,m}^n - T^\dagger y\right\| = \begin{cases} o(\delta^{2\nu/(2\nu+1)}) + o(\gamma_m^{2\nu}) & \text{for } \nu < 1/2, \\ o(\delta^{2\nu/(2\nu+1)}) + O(\gamma_m) & \text{for } 1/2 \leqslant \nu < n, \\ O(\delta^{2n/(2n+1)}) + O(\gamma_m) & \text{for } \nu \geqslant n. \end{cases}$$

Proof. It follows from Lemma 3.9 that there exists a number η, $C_1 \leqslant \eta \leqslant C_2$, such that

$$(1 - K(2n-1)b_m^2/\alpha_m(\delta)) f_n^m(\alpha_m(\delta), y) = \eta\delta^2.$$

Hence we obtain by Lemma 3.2, Lemma 3.7 and Lemma 3.10 that

$$\left\|x^n_{\alpha_m(\delta),\delta,m} - T^\dagger y\right\|^2 \leqslant 2/\min\{\eta,1\}\left(2n\eta\delta^2/\alpha_m(\delta) + \left\|x^n_{\alpha_m(\delta),m} - T^\dagger y\right\|^2\right)$$

$$\leqslant 2/\min\{\eta,1\}\min\Bigg\{2n\eta\delta^2/\alpha + \left\|x^n_{\alpha,m} - T^\dagger_m y\right\|^2 + Kb^2_m \sum_{j=1}^{2n-1} j\rho_{j,m}(\alpha)$$

$$+1/K\left\|(I-P_m)T^\dagger y\right\|^2 - 2\left\langle T^\dagger_m y, T^\dagger_m y - T^\dagger y\right\rangle$$

$$+\left\|T^\dagger_m y - T^\dagger y\right\|^2 : \alpha > 0\Bigg\}$$

$$\leqslant 2/\min\{C_1,1\}\inf\Bigg\{2nC_2\delta^2/\alpha + \left\|x^n_{\alpha,m} - T^\dagger y\right\|^2 + 2Kb^2_m \sum_{j=1}^{2n-1} j\rho_{j,m}(\alpha)$$

$$+2/K\left\|(I-P_m)T^\dagger y\right\|^2 : \alpha > 0\Bigg\}$$

$$\leqslant 2/\min\{C_1,1\}\Bigg(2nC_2\delta^2/\beta_{m,n}(\delta) + \left\|x^n_{\beta_{m,n}(\delta),m} - T^\dagger y\right\|^2$$

$$+2Kb^2_m \sum_{j=1}^{2n-1} j\rho_{j,m}(\beta_{m,n}(\delta)) + 2/K\left\|(I-P_m)T^\dagger y\right\|^2\Bigg),$$

where $\beta_{m,n}(\delta) := \max\{\beta_n(\delta), \gamma^2_m\}$, with $\beta_n(\delta)$ given by (2.9). Since

$$\left\|x^n_{\beta_{m,n}(\delta),m} - T^\dagger y\right\|^2 \leqslant \left(\left\|x^n_{\beta_{m,n}(\delta),m} - x^n_{\beta_{m,n}(\delta)}\right\| + \left\|x_{\beta_{m,n}(\delta)} - T^\dagger y\right\|\right)^2$$

$$\leqslant 2\left[\left\|x_{\beta_{m,n}(\delta),m} - x_{\beta_{m,n}(\delta)}\right\|^2 + \varphi_n(\beta_{m,n}(\delta))\right]$$

and $\beta_{m,n}(\delta) \geqslant \gamma^2_m$, we obtain by Lemma 3.5 and Lemma 3.8,

$$\left\|x^n_{\alpha_m(\delta),\delta,m} - T^\dagger y\right\|^2$$

$$\leqslant 2/\min\{C_1,1\}\Big(2nC_2\delta^2/\beta_{m,n}(\delta) + 2\varphi_n(\beta_{m,n}(\delta)) \tag{3.6}$$

$$+2g^2(\gamma_m) + 2Kh(\gamma_m) + 2/K\left\|(I-P_m)T^\dagger y\right\|^2\Big).$$

By the proof of Theorem 2.8 we have $\lim_{\delta\to 0}\beta_n(\delta) = 0$ and $\lim_{\delta\to 0}\delta^2/\beta_{m,n}(\delta) = 0$ because of $\beta_{m,n}(\delta) \geqslant \beta_n(\delta)$.

Further, our assumptions on V_m imply $\lim_{m\to\infty}\|(I-P_m)T^\dagger y\| = 0$ and $\lim_{m\to\infty}\gamma_m = 0$ (see, e.g., [12, Lemma 2.4.1]). This implies $\lim_{m\to\infty,\delta\to 0}\beta_{m,n}(\delta) = 0$ and hence $\lim_{m\to\infty,\delta\to 0}\varphi_n(\beta_{m,n}(\delta)) = 0$ by Lemma 2.1. Hence, all expressions on the right-hand side of (3.6) tend to zero, which proves convergence for our parameter choice.

Now assume $T^\dagger y \in R((T^*T)^\nu)$ for some $\nu > 0$. By Lemma 3.5, Theorem 3.6 and Lemma 3.8 we obtain

$$(3.7)\quad 2g^2(\gamma_m) + 2Kh(\gamma_m) + (2/K)\|(I - P_m)T^\dagger y\|^2 = \begin{cases} o(\gamma_m^{4\nu}) & \text{for } \nu < 1/2, \\ O(\gamma_m^2) & \text{for } \nu \geqslant 1/2. \end{cases}$$

Since $\beta_n(\delta) \leqslant \beta_{m,n}(\delta)$, we obtain by the proof of Theorem 2.8,

$$(3.8)\qquad 2n\delta^2/\beta_{m,n}(\delta) \leqslant \varphi_n(\beta_n(\delta)) = \begin{cases} o(\delta^{4\nu/(2\nu+1)}) & \text{for } \nu < n, \\ O(\delta^{4n/(2n+1)}) & \text{for } \nu \geqslant n. \end{cases}$$

Further, the results of Schock [19] imply that

$$(3.9)\qquad \varphi_n(\gamma_m^2) = \begin{cases} o(\gamma_m^{4\nu}) & \text{for } \nu < n, \\ O(\gamma_m^{4n}) & \text{for } \nu \geqslant n. \end{cases}$$

Hence, by (3.6), (3.7), (3.8) and (3.9) we obtain

$$\left\| x^n_{\alpha_m(\delta),\delta,m} - T^\dagger y \right\|^2 = \begin{cases} o(\delta^{4\nu/(2\nu+1)}) + o(\gamma_m^{4\nu}) & \text{for } 0 < \nu < 1/2, \\ o(\delta^{4\nu/(2\nu+1)}) + O(\gamma_m^2) & \text{for } 1/2 \leqslant \nu < n, \\ O(\delta^{4n/(2n+1)}) + O(\gamma_m^2) & \text{for } \nu \geqslant n, \end{cases}$$

which completes our proof. □

Remark 3.12. Although convergence rates are expressed in terms of γ_m, only the numbers b_m are used for the computation of $\alpha_m(\delta)$. Note that the numbers b_m are effectively computable (cf. Appendix 3), whereas only estimates for γ_m are available in general. Groetsch [12], resp. Engl and Neubauer [6], use information about γ_m to compute the regularization parameter. This might be detrimental for actual computations, since poor estimates for γ_m could also yield poor convergence rates.

Remark 3.13. For actual computations the choice of K, C and n is of course important. In view of the proof of Lemma 3.7, it seems to be advantageous to choose K small, if one has the a priori information that $T^\dagger y$ may be approximated well by elements of V_m, i.e., $\|(I - P_m)T^\dagger y\|$ is small. More critical is the choice of C and n. In the infinite-dimensional case, a choice of $C = 1$ is optimal by Remark 2.4, and the approximation error remains uniformly bounded for all n by Theorem 2.9. Our numerical experience suggests that analogous results could also be expected for our finite-dimensional approximations, but we were not able to prove this.

4. Numerical Results. All examples are Fredholm integral equations of the first kind on $[0,1]$,

$$\int_0^1 k(t,s)x(s)\,ds = y(t),$$

where the kernel is given by

$$k(t,s) := \begin{cases} t(1-s) & \text{if } t \leqslant s, \\ s(1-t) & \text{if } t > s. \end{cases}$$

This kernel is the Green's function of the vibrating string with fixed ends. It is well known (cf. [2]) that $\{u_j;\ v_j;\ \sigma_j\}$ with $u_j(t) = v_j(t) = 2^{1/2}\sin(j\pi t)$, $t \in [0,1]$, and $\sigma_j = (\pi j)^{-2}$ forms a singular system for this operator.

In our examples, we chose V_m as a space of linear splines on a uniform grid of $(m + 1)$ points in $[0, 1]$. For some computational aspects we refer to Appendix 3 in the Supplements section. We obtained the following results for b_m:

m	4	8	16	32
b_m	0.548×10^{-2}	0.131×10^{-2}	0.325×10^{-3}	0.811×10^{-4}

It appears that $b_m = O(m^{-2})$, which agrees with the fact that $\gamma_m = O(m^{-2})$ (cf. [9]).

Example 4.1. (a) $y(t) = (t - t^3)/6$, $T^\dagger y(s) = s$. Since $\int_0^1 s \sin(j\pi s)\,ds = (-1)^{j+1}/j\pi$, we obtain that $T^\dagger y \in R((T^*T)^\nu)$ for $\nu < 1/8$. (Note that $T^\dagger y \in R((T^*T)^\nu)$ if and only if $\sum_j \sigma_j^{-4\nu}\langle T^\dagger y, u_j\rangle^2 < \infty$.) Hence, according to theory, we should obtain a convergence rate $o(\gamma_m^{2\nu})$ for $\nu < 1/8$, independent of the iteration number n. For $K = 1$, the results were as follows:

	$n = 1$		$n = 10$		$n = 100$	
m	e_m	$e_m b_m^{-1/4}$	e_m	$e_m b_m^{-1/4}$	e_m	$e_m b_m^{-1/4}$
4	0.167	0.614	0.153	0.562	0.152	0.559
8	0.114	0.600	0.104	0.547	0.103	0.541
16	0.806×10^{-1}	0.600	0.731×10^{-1}	0.544	0.723×10^{-1}	0.538
32	0.570×10^{-1}	0.600	0.517×10^{-1}	0.544	0.511×10^{-1}	0.538

Here, $e_m := \|x^n_{\alpha_m, m} - T^\dagger y\|$ with $\alpha_m := (2n - 1)Kb_m^2$. The columns headed by $e_m b_m^{-1/4}$ show that the convergence seems to be $O(b_m^{1/4})$.

Since $T^\dagger y \in V_m$ for each m, we should obtain better results for small K, according to Remark 3.13. The following table shows that for $K = 0.05$ the absolute errors are significantly smaller. However, the convergence rates seem to be as before.

	$n = 1$		$n = 10$		$n = 100$	
m	e_m	$10 \cdot e_m \cdot b_m^{-1/4}$	e_m	$10^3 \cdot e_m b_m^{-1/4}$	e_m	$10^4 \cdot e_m b_m^{-1/4}$
4	0.252×10^{-1}	0.926	0.947×10^{-3}	3.48	0.245×10^{-3}	9.00
8	0.159×10^{-1}	0.836	0.221×10^{-3}	1.16	0.218×10^{-4}	1.15
16	0.111×10^{-1}	0.830	0.145×10^{-3}	1.08	0.131×10^{-4}	0.976
32	0.784×10^{-2}	0.826	0.102×10^{-3}	1.07	0.925×10^{-5}	0.975

(b) y as in (a); for each m, y was 30 times randomly perturbed with $\delta_m = b_m/2$. Choosing the regularization parameter as the solution of (3.4), the convergence rate should be $o(b_m^\mu)$ with $\mu < 0.2$. The following results were computed with $C = 1.01$ and $K = 0.01$.

	$n = 1$		$n = 10$		$n = 100$	
m	$\tilde{e}_m$	$\tilde{e}_m b_m^{-0.2}$	$\tilde{e}_m$	$\tilde{e}_m b_m^{-0.2}$	$\tilde{e}_m$	$\tilde{e}_m b_m^{-0.2}$
4	0.336	0.952	0.338	0.958	0.339	0.960
8	0.242	0.913	0.235	0.886	0.235	0.886
16	0.176	0.877	0.175	0.872	0.175	0.872
32	0.131	0.862	0.131	0.862	0.130	0.855

Here and below, $\tilde{e}_m$ denotes the maximum error of all tests.

Example 4.2. (a) $y(t) = (1/24)(t - 2t^3 + t^4)$, $T^\dagger y(s) = (1/2)(s - s^2)$. Since $\int_0^1 T^\dagger y(s)\,\sin(j\pi s)\,ds = (-1)^j/(\pi j)^3$, we obtain $T^\dagger y \in R((T^*T)^\nu)$ for $\nu < 5/8$. Hence the convergence rate should be $O(\gamma_m)$. For $K = 1$, we obtain the following results:

	$n = 1$		$n = 2$		$n = 4$	
m	e_m	$e_m b_m^{-1}$	e_m	$e_m b_m^{-1}$	e_m	$e_m b_m^{-1}$
4	0.250×10^{-2}	0.456	0.242×10^{-2}	0.442	0.240×10^{-2}	0.438
8	0.601×10^{-3}	0.458	0.594×10^{-3}	0.453	0.592×10^{-3}	0.452
16	0.149×10^{-3}	0.455	0.147×10^{-3}	0.452	0.147×10^{-3}	0.452
32	0.367×10^{-4}	0.453	0.366×10^{-4}	0.451	0.365×10^{-4}	0.450

The following table gives the errors for the best possible approximation of $T^\dagger y$ by elements of V_m.

m	4	8	16	32
$\|(I - P_m)T^\dagger y\|$	0.233×10^{-2}	0.582×10^{-3}	0.146×10^{-3}	0.364×10^{-4}

(b) y as in (a); y was 30 times randomly perturbed with $\delta_m := b_m$ for each m. According to theory, the convergence rate should be $o(b_m^\mu)$ with $\mu < 5/9$.

	$n = 1$		$n = 2$		$n = 4$	
m	$\tilde{e}_m$	$\tilde{e}_m b_m^{-5/9}$	$\tilde{e}_m$	$\tilde{e}_m b_m^{-5/9}$	$\tilde{e}_m$	$\tilde{e}_m b_m^{-5/9}$
4	0.770×10^{-1}	1.39	0.746×10^{-1}	1.35	0.731×10^{-1}	1.32
8	0.272×10^{-1}	1.09	0.224×10^{-1}	0.895	0.200×10^{-1}	0.799
16	0.971×10^{-2}	0.842	0.728×10^{-2}	0.631	0.601×10^{-2}	0.521
32	0.451×10^{-2}	0.845	0.372×10^{-2}	0.697	0.344×10^{-2}	0.644

Example 4.3. (a) $y(t) = (1/30)(3t - 5t^3 + 3t^5 - t^6)$, $T^\dagger y(s) = s - 2s^3 + s^4$. In this example one has $T^\dagger y \in R((T^*T)^\nu)$ for $\nu < 9/8$. Hence, the convergence rate should again be $O(\gamma_m)$. If we choose $K = 1$, we obtain the following results:

	$n = 1$		$n = 2$		$n = 4$	
m	e_m	$e_m b_m^{-1}$	e_m	$e_m b_m^{-1}$	e_m	$e_m b_m^{-1}$
4	0.536×10^{-2}	0.978	0.531×10^{-2}	0.969	0.531×10^{-2}	0.969
8	0.129×10^{-2}	0.985	0.129×10^{-2}	0.985	0.129×10^{-2}	0.985
16	0.320×10^{-3}	0.985	0.320×10^{-3}	0.985	0.320×10^{-3}	0.985
32	0.798×10^{-4}	0.984	0.798×10^{-4}	0.984	0.798×10^{-4}	0.984

Note that our parameter choice yields nearly the best possible approximation $P_m T^\dagger y$:

m	4	8	16	32
$\|(I - P_m)T^\dagger y\|$	0.531×10^{-2}	0.129×10^{-2}	0.320×10^{-3}	0.798×10^{-4}

(b) y as in (a); y was again 30 times randomly perturbed with $\delta_m := b_m$ for each m. In this case we should obtain, for $n = 1$, convergence rate $O(b_m^\mu)$ with $\mu = 2/3$, resp., for $n \geq 2$, the convergence rate $o(b_m^\mu)$ with $\mu < 9/13$. A choice $C = 1.01$, $K = 1$ yields:

m	$n = 1$		$n = 2$		$n = 4$	
	$\tilde{e}_m$	$\tilde{e}_m b_m^{-2/3}$	e_m	$\tilde{e}_m b_m^{-9/13}$	$\tilde{e}_m$	$\tilde{e}_m b_m^{-9/13}$
4	0.124	3.99	0.112	4.12	0.106	3.90
8	0.353×10^{-1}	2.95	0.249×10^{-1}	2.47	0.201×10^{-1}	1.99
16	0.116×10^{-1}	2.45	0.662×10^{-2}	1.72	0.474×10^{-2}	1.23
32	0.467×10^{-2}	2.49	0.236×10^{-2}	1.60	0.165×10^{-2}	1.12

Example 4.4. $y(t) = (1/\pi^2)\sin(\pi t)$, $T^\dagger y(s) = \sin \pi s$. In this example, $T^\dagger y \in R((T^*T)^\nu)$ holds for any $\nu > 0$. If the data are known exactly, we again obtain nearly the best possible approximation $P_m T^\dagger y$. If the data are randomly perturbed with $\delta_m = b_m$, we should obtain, for each $n \in \mathbf{N}$, the convergence rate $O(b_m^{2n/(2n+1)})$. The test was performed 30 times again; the following table shows the maximum error observed for the choice $C = 1.01$, $K = 1$.

m	$n = 1$		$n = 2$		$n = 4$	
	$\tilde{e}_m$	$\tilde{e}_m b_m^{-2/3}$	$\tilde{e}_m$	$\tilde{e}_m b_m^{-4/5}$	$\tilde{e}_m$	$\tilde{e}_m b_m^{-8/9}$
4	0.170	5.47	0.135	8.70	0.117	12.0
8	0.540×10^{-1}	4.51	0.336×10^{-1}	6.80	0.248×10^{-1}	9.05
16	0.178×10^{-1}	3.77	0.875×10^{-2}	5.40	0.572×10^{-2}	7.21
32	0.693×10^{-2}	3.71	0.292×10^{-2}	5.47	0.169×10^{-2}	7.31

Johannes-Kepler-Universität
A-4040 Linz, Austria

1. R. ARCANGELI, "Pseudo-solution de l'équation $Ax = y$," *C. R. Acad. Sci. Paris Sér. A*, v. 263, 1966, pp. 282–285.

2. R. COURANT & D. HILBERT, *Methoden der Mathematischen Physik I*, Springer, Berlin, 1968.

3. H. W. ENGL, "Necessary and sufficient conditions for convergence of regularization methods for solving linear operator equations of the first kind," *Numer. Funct. Anal. Optim.*, v. 3, 1981, pp. 201–222.

4. H. W. ENGL, "Discrepancy principles for Tikhonov regularization of ill-posed problems leading to optimal convergence rates," *J. Optim. Theory Appl.*, v. 52, 1987, pp. 209–215.

5. H. W. ENGL, "On the choice of the regularization parameter for iterated Tikhonov-regularization of ill-posed problems," *J. Approx. Theory*, v. 49, 1987, pp. 55–63.

6. H. W. ENGL & A. NEUBAUER, "A variant of Marti's method for solving ill-posed linear integral equations that leads to optimal convergence rates," *Math. Comp.*, v. 45, 1985, pp. 405–416.

7. G. H. GOLUB & C. F. VAN LOAN, *Matrix Computations*, The Johns Hopkins University Press, Baltimore, MD, 1983.

8. C. W. GROETSCH, *Elements of Applicable Functional Analysis*, Marcel Dekker, New York, 1980.

9. C. W. GROETSCH, J. T. KING & D. MURIO, "Asymptotic analysis of a finite element method for Fredholm equations of the first kind," in *Treatment of Integral Equations by Numerical Methods* (C. T. H. Baker and G. F. Miller, eds.), Academic Press, London, 1982.

10. C. W. GROETSCH, "Comments on Morozov's discrepancy principle," in *Improperly Posed Problems and Their Numerical Treatment* (G. Hämmerlin and K. H. Hoffmann, eds.), Birkhäuser, Basel, 1983.

11. C. W. GROETSCH, "On the asymptotic order of accuracy of Tikhonov regularization," *J. Optim. Theory Appl.*, v. 41, 1983, pp. 293–298.

12. C. W. GROETSCH, *The Theory of Tikhonov Regularization for Fredholm Equations of the First Kind*, Pitman, Boston, 1984.

13. C. W. GROETSCH & E. SCHOCK, "Asymptotic convergence rate of Arcangeli's method for ill-posed problems," *Applicable Anal.*, v. 18, 1984, pp. 175–182.

14. J. T. KING & D. CHILLINGWORTH, "Approximation of generalized inverses by iterated regularization," *Numer. Funct. Anal. Optim.*, v. 2, 1979, pp. 449–513.

15. J. T. MARTI, "An algorithm for computing minimum norm solutions of Fredhom integral equations of the first kind," *SIAM J. Numer. Anal.*, v. 15, 1978, pp. 1071–1076.

16. A. MOROZOV, "On the solution of functional equations by the method of regularization," *Soviet Math. Dokl.*, v. 7, 1966, pp. 414–417.

17. M. Z. NASHED (Editor), *Generalized Inverses and Applications*, Academic Press, New York, 1976.

18. E. SCHOCK, "Approximate solution of ill-posed equations: arbitrarily slow convergence vs. superconvergence," in *Improperly Posed Problems and Their Numerical Treatment* (G. Hämmerlin and K. H. Hofmann, eds.), Birkhäuser, Basel, 1983.

19. E. SCHOCK, "On the asymptotic order of accuracy of Tikhonov regularization," *J. Optim. Theory Appl.*, v. 44, 1984, pp. 95–104.

20. E. SCHOCK, "Parameter choice by discrepancy principles for the Tikhonov regularization of ill-posed problems," *Integral Equations Operator Theory*, v. 7, 1984, pp. 895–898.

21. T. I. SEIDMAN, "Nonconvergence results for the application of least-squares estimation to ill-posed problems," *J. Optim. Theory Appl.*, v. 30, 1980, pp. 535–547.

22. J. STOER & R. BULIRSCH, *Introduction to Numerical Analysis*, Springer, New York, 1980.