An Improved Version of Marti’s Method
for Solving Ill-Posed Linear Integral Equations

By Heinz W. Engl and Andreas Neubauer*

Abstract. We propose an algorithm for solving linear integral equations of the first kind that can be viewed as a variant of Marti’s method; as opposed to that method, our algorithm leads to optimal convergence rates (also with noisy data).

1. Introduction. Throughout this paper, let $X$ and $Y$ be real Hilbert spaces, $T: X \to Y$ a bounded linear operator with nonclosed range. Then the problem of determining the “best-approximate solution” of

$$T x = y$$

is ill-posed: The best-approximate solution exists only for $y \in D(T^+) = R(T) + R(T)^\perp$ (which we assume from now on) and depends discontinuously on the right-hand side. Here $T^+$ is the Moore-Penrose inverse of $T$ (see [12]); the best-approximate solution is defined as the element of minimal norm that minimizes the residual $\|Tx - y\|$ and can be written as $T^+y$. A prominent example for the ill-posed case of (1.1) is a Fredholm integral equation of the first kind. Ill-posed problems have to be solved by regularization methods, e.g., Tikhonov regularization. See [7] and [14] for more background.

An algorithm that has been used successfully in recent years is “Marti’s method” (see [9], [10], [11]). In this method, a sequence of finite-dimensional subspaces $V_1 \subseteq V_2 \subseteq V_3 \subseteq \cdots$ of $X$ with $\bigcup_{n \in \mathbb{N}} V_n = X$ is used to compute approximate solutions of (1.1) as follows:

Let, for all $n \in \mathbb{N}$,

$$a_n := \inf \{ \|Tx - y\| / x \in V_n \},$$

$P_n$ be the orthogonal projector of $X$ onto $V_n$, and $b_n > 0$ be chosen such that

$$\lim_{n \to \infty} b_n = 0, \quad \lim_{n \to \infty} b_n = 0$$

holds. Then $x_n$ is defined by

$$x_n \in V_n,$$

$$\|Tx_n - y\|^2 \leq a_n^2 + b_n^2,$$
\[
\|x_n\| = \inf\{\|x\|/x \in V_n \text{ and } x \text{ fulfills (1.5)}\}.
\]

This is Marti’s algorithm as given in [11]; originally ([9], [10]), the right-hand side of (1.5) was \((a_n + b_n)^2\) and the algorithm was formulated only for \(y \in R(T)\). The remarks of this section apply to both versions of the algorithm.

Marti proved that \(\{x_n\}\) converges to \(T^*y\) as \(n \to \infty\) and that
\[
\|x_n - T^*y\| = O(\|b_n\|)
\]
if
\[
T^*y \in R(T^*)
\]
holds, which can be interpreted as an a priori smoothness assumption about the (unknown) exact solution. By \(T^*\) we denote the adjoint of \(T\).

Incidentally, in [8] it is claimed that Marti’s results are wrong as stated; this statement is formally correct if applied to [9], since there the condition (1.3) is missing. However, this condition appears in [10], and the results in that paper are correct. It has to be pointed out that the authors of [8] quote [10], so that their claim that Marti’s result is in error and that they corrected it is not justified, since Marti himself had corrected his error before them in [10]. Moreover, it is easy to see where the error in [9] is: There, (5) is wrong, since (in the notation of that paper) \(P_{m_0}f_0\) need not be in \(V_n\). From there, one can immediately deduce the condition that is added in the corrected version of Marti’s result ([8, Theorem 2.5]).

In [7, Section 4.3], C. W. Groetsch has given the following alternate formulation of Marti’s algorithm:

\(x_n\) is determined by
\[
\begin{align*}
\alpha_n x_n + T_n^* T_n x_n &= T_n^* y, \\
\|T x_n - y\|^2 &= a_n^2 + b_n^2,
\end{align*}
\]
where
\[
T_n := T|_{V_n}.
\]
Note that (1.9) is just Tikhonov regularization, where the regularization parameter \(\alpha_n\) is determined from (1.10), which can be interpreted as a “discrepancy principle” (cf. [13], [7, Section 3.3]). In view of [5] it cannot be expected that the convergence rate in (1.7) can be improved, even under stronger smoothness assumptions. However, our results in [2] and [3] can be used to modify (1.9), (1.10) (and thus Marti’s method) in such a way that the convergence rate in (1.7) is improved to the best rate one can reasonably expect.

2. Main Results. As in Marti’s method, let \(V_1 \subseteq V_2 \subseteq V_3 \subseteq \cdots\) be a sequence of finite-dimensional subspaces of \(X\) with \(\bigcup_{n \in \mathbb{N}} V_n = X\). For each \(n \in \mathbb{N}\), let
\[
\gamma_n := \|(I - P_n) T^*\|,
\]
where \(P_n\) is the orthogonal projector of \(X\) onto \(V_n\), and let \(y_n \in Y\) with
\[
\|y - y_n\| \leqslant \delta_n,
\]
where the sequence \(\{\delta_n\}\) is assumed to be known. Let \(\{b_n\}\) be a sequence in \(\mathbb{R}^+\); we assume
\[
\lim_{n \to \infty} b_n = 0, \quad \lim_{n \to \infty} \delta_n = 0.
\]
Further properties of \( \{b_n\} \) will be fixed below. For each \( \alpha > 0 \) and \( n \in \mathbb{N} \) we denote by \( x^\delta_{n,\alpha} \) the unique solution of
\begin{equation}
\alpha x + T^*_n T_n x = T^*_n y_n
\end{equation}
in \( V_n \), where \( T_n \) is as in (1.11); with this definition, let
\begin{equation}
\rho_n(\alpha) := \left\| T^*_n T_n x^\delta_{n,\alpha} - T^*_n y_n \right\|^2.
\end{equation}
Note that \( \rho_n \) depends also on \( y_n \) and hence on \( \delta_n \), which is the norm of the data error.

In our variant of Marti's method, the data error will be included from the beginning, since despite its different origin, it can be treated the same way as the approximation error symbolized by \( b_n \).

Let \( D_1, D_2 \) be positive constants, and \( p, q > 0 \) be parameters that are fixed later. In our algorithm, \( x_n \in V_n \) is defined as the unique solution of (2.4), where \( \alpha \) is the solution of
\begin{equation}
\rho_n(\alpha) = (D_1 b_n + D_2 \delta_n)^p \alpha^{-q}.
\end{equation}
This can be seen as a variant of Marti's method as formulated by (1.9), (1.10) in the following sense:

The equations (1.9) and (2.4) are identical except that in (2.4) we only use the approximation \( y_n \) for \( y \); as can be seen from [7, (4.3.5)], (1.10) can also be written as
\begin{equation}
\left\| T_n x_n - Q_n y \right\|^2 = b_n^2,
\end{equation}
where \( Q_n \) is the orthogonal projector of \( Y \) onto \( T(V_n) \). If we replace the residual in (2.7) by the residual of the corresponding finite-dimensional normal equation, we obtain (2.6) with \( \delta_n = 0, D_1 = 1, p = 2, \) and \( q = 0 \).

From now on we denote by \( x^\delta_n \) and \( x_n \) always the unique solution in \( V_n \) of (2.4) and (2.4) with \( y_n \) replaced by \( y \), respectively, where \( \alpha \) is determined by (2.6); this \( \alpha \) will also be denoted by \( \alpha_n \). Of course, these quantities depend on \( p \) and \( q \). We will determine \( p \) and \( q \) in such a way that the convergence rate \( \|x^\delta_n - T^*y\| \) is best possible.

To exclude trivial cases, we will always assume that
\begin{equation}
T^*y \neq 0 \quad \text{and} \quad T^*_n y_n \neq 0,
\end{equation}
where the latter assumption follows from the first when \( n \) is sufficiently large. The first result shows that our algorithm makes sense:

**Proposition 2.1.** For any \( p, q > 0 \), (2.6) is uniquely solvable.

**Proof.** It follows from [4, Lemma 3.1] (cf. also [7, Theorem 3.3.1]), applied with \( T_n, V_n \) and \( y_n \) instead of \( T, X \) and \( b^\delta \), respectively, that \( \rho_n \) is continuous, strictly increasing and that
\[
\lim_{\alpha \to 0} \rho_n(\alpha) = 0 \quad \text{and} \quad \lim_{\alpha \to +\infty} \rho_n(\alpha) = \left\| T^*_n y_n \right\|^2 > 0
\]
because of (2.8). Thus, the assertion follows from the Intermediate Value Theorem. \( \square \)

Since \( x^\delta_n \) and \( x_n \) minimize the Tikhonov functionals
\[
x \to \|Tx - y_n\|^2 + \alpha_n \|x\|^2 \quad \text{and} \quad x \to \|Tx - y\|^2 + \alpha_n \|x\|^2,
\]
respectively, over the subspace \( V_n \), we obtain from [7, Lemma 4.2.7] that
\[
\|x_n^\delta - x_n\| \leq \delta_n \cdot \alpha_n^{-1/2}.
\]
For the same reason, we can use [7, Lemma 4.2.3] to obtain
\[
\|x_n - x_\alpha\| \leq \left(1 + \frac{\gamma_n^2}{\alpha_n}\right)^{1/2} \cdot \|(I - P_n)x_\alpha\|,
\]
where \( \gamma_n \) is defined in (2.1), and where
\[
x_\alpha := (T^*T + \alpha I)^{-1}T*y
\]
is the approximate solution obtained from infinite-dimensional Tikhonov regularization; of course, \( x_\alpha \) is not used in our algorithm, which is strictly finite-dimensional.

**Lemma 2.2.** \( \lim_{n \to \infty} \alpha_n = 0 \).

**Proof.** One shows, as in the proof of [2, Lemma 2.2], that if \( \{\alpha_n\} \) would have a subsequence (again denoted by \( \{\alpha_n\} \)) with
\[
\lim_{n \to \infty} \alpha_n = +\infty,
\]
then this would imply
\[
\lim_{n \to \infty} x_n^\delta = 0,
\]
and hence
\[
\lim_{n \to \infty} \rho_n(\alpha_n) = \|T*y\|^2
\]
because of (2.2), (2.3) and the facts that \( T_n^* = P_nT^* \) and \( P_n \to I \) pointwise. Hence
\[
0 = \lim_{n \to \infty} (D_1b_n + D_2\delta_n) = \lim_{n \to \infty} (\alpha_n^q \cdot \rho_n(\alpha_n)) = +\infty
\]
because of (2.12), (2.14), and (2.8), which is a contradiction. Thus, (2.12) cannot hold for any subsequence, which implies the boundedness of \( \{\alpha_n\} \).

Now assume that
\[
\lim_{n \to \infty} \alpha_n = \alpha > 0
\]
holds for some subsequence (again denoted by \( \{\alpha_n\} \)). Then
\[
\lim_{n \to \infty} \left\|(T_n^*T_n + \alpha_n I)^{-1}T_n^* - (T_n^*T_n + \alpha I)^{-1}T_n^*\right\| = 0,
\]
so that
\[
\lim_{n \to \infty} \|x_n^\delta - (T_n^*T_n + \alpha I)^{-1}T_n*y\| = 0.
\]
Because of [7, Lemma 4.2.3],
\[
\left\|(T_n^*T_n + \alpha I)^{-1}T_n*y - x_\alpha\right\|^2 \leq \left(1 + \frac{\gamma_n^2}{\alpha}\right) \cdot \|(I - P_n)x_\alpha\|^2,
\]
where \( x_\alpha \) is as in (2.11) with \( \alpha \) as in (2.15). Since the right-hand side of (2.17) tends to 0 as \( n \to \infty \), (2.16) and (2.17) imply
\[
\lim_{n \to \infty} \|x_n^\delta - x_\alpha\| = 0.
\]
Since, as easy calculations show,
\begin{equation}
\label{eq:2.19}
\rho_n(\alpha) = \alpha^2 \cdot \|x_n^{\delta_n}\|^2
\end{equation}
holds for all $\alpha > 0$, (2.18) implies together with the definition of $\alpha_n$ that
\begin{equation}
0 = \lim_{n \to \infty} (D_1 b_n + D_2 \delta_n)^p = \lim_{n \to \infty} \left( \alpha_n^{p-2} \|x_n^{\delta_n}\|^2 \right) = \alpha_n^{p-2} \|x^\star\|^2,
\end{equation}
which implies $x^\star = 0$ in contradiction to (2.8). Hence (2.15) cannot hold for any subsequence, which proves the lemma. \hfill \Box

**Lemma 2.3.** If $0 < p \leq 2q$, then $\lim_{n \to \infty} (\delta_n^2 \cdot \alpha_n^{-1}) = 0$.

**Proof.** Since $x_n^{\delta_n}$ minimizes the functional $x \to \|Tx - Q_n y_n\|^2 + \alpha_n \|x\|^2$ over $V_n$, where $Q_n$ is the orthogonal projector of $Y$ onto $T(V_n)$ (and hence $T^*Q_n = T^*$), we have
\begin{equation}
0 < \rho_n(\alpha_n) \leq \|T_n^*\|^2 \cdot \left( \|T_n x_n^{\delta_n} - Q_n y_n\|^2 + \alpha_n \|x_n^{\delta_n}\|^2 \right)
\leq \|T\|^2 \cdot \left( \|TP_n T^\dagger y - Q_n y_n\|^2 + \alpha_n \|T^\dagger y\|^2 \right).
\end{equation}
Since $P_n T^\dagger y \to T^\dagger y$ and $Q_n$ converges pointwise to the projector of $Y$ onto $R(T)$, i.e., to the extension of $TT^\dagger$ onto $Y$, $\|TP_n T^\dagger y - Q_n y_n\| \to 0$. This implies, together with the last inequality and with Lemma 2.2, that $\lim_{n \to \infty} \rho_n(\alpha_n) = 0$. Hence,
\begin{equation}
0 \leq \delta_n^p \cdot \alpha_n^{-p/2} \leq D_2^p \cdot (D_1 b_n + D_2 \delta_n)^p \cdot \alpha_n^{-p/2} = D_2^p \cdot \alpha_n^{-p/2} \cdot \rho_n(\alpha_n);
\end{equation}
the last expression tends to 0 as $n \to \infty$ because of the first part of this proof, Lemma 2.2 and the assumptions on $p$ and $q$. Hence $\lim_{n \to \infty} \delta_n^p \cdot \alpha_n^{-p/2} = 0$, which implies the assertion. \hfill \Box

The next result, where a condition on $\{b_n\}$—the analogue of (1.3) in Marti’s method—is introduced, shows that for a wide range of $p$ and $q$, our method converges.

**Theorem 2.4.** Let $0 < p \leq 2q$ and assume that
\begin{equation}
\label{eq:2.20}
\gamma_n = O(b_n^{p/2q}),
\end{equation}
where $\gamma_n$ is defined in (2.1). Then $\lim_{n \to \infty} x_n^{\delta_n} = T^\dagger y$.

**Proof.** We first show that
\begin{equation}
\label{eq:2.21}
\{ \gamma_n^2/\alpha_n \} \text{ is bounded.}
\end{equation}
To see this, note that (cf. the proof of Proposition 2.1)
\begin{equation}
(D_1 b_n + D_2 \delta_n)^p \cdot \alpha_n^{-q} = \rho_n(\alpha_n) \leq \|T_n^* y_n\|^2 \cdot (\|y\|^2 + \delta_n)^2,
\end{equation}
hence $\|x_n^{\delta_n} - T^\dagger y\|^2 \leq \|x_n^{\delta_n} - x_n\|^2 + \|x_n - x_n^{\alpha_n}\|^2 + \|x_n^{\alpha_n} - T^\dagger y\|^2$,
where $x_n^{\alpha_n}$ is defined by (2.11) for $\alpha = \alpha_n$. Because of (2.9) and Lemma 2.3, the first term of this estimate tends to 0. Because of Lemma 2.2 and standard results about infinite-dimensional Tikhonov regularization (see, e.g., [7, Sections 2.1 and 3.1]), the third term tends to 0. Because of (2.10), the second term can be estimated as follows:
\begin{equation}
\|x_n - x_n^{\alpha_n}\|^2 \leq \left( 1 + \frac{\gamma_n^2}{\alpha_n} \right)^{1/2} \cdot \left[ \|(I - P_n)\| \cdot \|x_n^{\alpha_n} - T^\dagger y\| + \|(I - P_n)T^\dagger y\| \right].
\end{equation}
which tends to 0 because of (2.21), the convergence of \( x^\alpha_n \) to \( T^\dagger y \) and the fact that \( \{ P_n \} \to I \) pointwise. This implies that \( \{ x^\delta_n \} \to T^\dagger y \). \( \square \)

Now we proceed to give estimates for the rate of convergence for various choices of \( p \) and \( q \). For this, we need the following estimate:

**Lemma 2.5.** Let the assumptions of Theorem 2.4 be fulfilled. Then there exist constants \( C_1, C_2 > 0 \) such that for all integers \( n \), \( C_1 < \langle D_{x_n} + D_2 \delta_n \rangle_{\alpha_n - q}^{-2} \leq C_2 \) holds.

**Proof.** Because of the definition of \( \alpha_n \) and (2.19), \( \langle D_{x_n} + D_2 \delta_n \rangle_{\alpha_n - q}^{-2} = \| x^\delta_n \|^2 \), which converges to \( \| T^\dagger y \|^2 \). Together with (2.8), this implies the assertion. \( \square \)

As usual with ill-posed problems, convergence rates can only be obtained under a priori assumptions about the exact solution:

**Theorem 2.6.** (a) Assume that (1.8) holds, and let \( p - 2 = q \geq 2 \) and

\[ \gamma_n = O\left( b_n^{1/2 + 1/q} \right). \]

Then \( \| x^\delta_n - T^\dagger y \| = O((b_n + \delta_n)^{1/2}) \).

(b) Assume that

\[ T^\dagger y \in R(T^\ast T) \]

and let \( \frac{1}{2}p - 2 = q \geq 1 \) and

\[ \gamma_n = \begin{cases} O\left( b_n^{1/3 + 2/3q} \right) & \text{if } q < 2, \\ O\left( b_n^{2/3} \right) & \text{if } q \geq 2. \end{cases} \]

Then \( \| x^\delta_n - T^\dagger y \| = O((b_n + \delta_n)^{2/3}) \).

**Proof.** Because of (2.9), \( \| x^\delta_n - x_n \| \leq D_2^{-1} \cdot (D_1 b_n + D_2 \delta_n) \cdot \alpha^{-1/2}_n \). Since (2.22) as well as (2.24) imply (2.20), we can apply Lemma 2.5 and obtain with \( C_1 \) and \( C_2 \) as defined there:

\[ \| x^\delta_n - x_n \| \leq C_3 \cdot (D_1 b_n + D_2 \delta_n)^{1 - p/2(q + 2)}, \quad \text{where } C_3 := D_2^{-1} \cdot C_2^{1/2(q + 2)}. \]

Hence

\[ \| x^\delta_n - T^\dagger y \| \leq \| x_n - T^\dagger y \| + C_3 \cdot (D_1 b_n + D_2 \delta_n)^{1 - p/2(q + 2)}. \]

Now, assume \( p - 2 = q \geq 2 \) and (2.22) (and hence \( p < 2q \) and (2.20)). Because of Lemma 2.5, \( (D_1 b_n + D_2 \delta_n)^{1/2} = O(\alpha_n^{q+1/2}) \). Hence

\[ \gamma_n = O\left( b_n^{1/2 + 1/q} \right) = O\left( b_n^{1/2} \right) = O\left( (D_1 b_n + D_2 \delta_n)^{1/2} \right) = O\left( \alpha_n^{1/2} \right). \]

Thus, we can apply [7, Theorem 4.2.5] and obtain

\[ \| x_n - T^\dagger y \| = O\left( \alpha_n^{1/2} \right). \]

Because of Lemma 2.5,

\[ \alpha_n^{1/2} = O\left( (D_1 b_n + D_2 \delta_n)^{p/2(q + 2)} \right) = O\left( (D_1 b_n + D_2 \delta_n)^{1/2} \right), \]
so that (2.25) and (2.26) imply
\[ \|x_n^{\delta_n} - T^\dagger y\| = O\left( (D_1 b_n + D_2 \delta_n)^{1/2} \right) = O\left( (b_n + \delta_n)^{1/2} \right). \]

This proves (a).

Now, let the assumptions of part (b) hold. Since also \( p \leq 2q \) and (2.20) hold, we can apply Lemma 2.3 and obtain
\[ (2.27) \quad (D_1 b_n + D_2 \delta_n)^{2/3} = O\left( \alpha_n \right) \]
and
\[ (2.28) \quad \alpha_n = O\left( (D_1 b_n + D_2 \delta_n)^{2/3} \right). \]

For \( q > 2 \), (2.24) and (2.27) imply \( \gamma_n = O(\alpha_n) \). For \( q < 2 \), we obtain the same result since \( b_n^{2/3 + 2/3q} = O(b_n^{2/3}) \). Thus we can apply [7, Theorem 4.2.6] to obtain
\[ (2.29) \quad \|x_n - T^\dagger y\| = O\left( \alpha_n \right). \]

The result of part (b) now follows from (2.25), (2.27), (2.29), and (2.28). □

**Remark 2.7.** Theorem 2.6 shows that our method improves Marti’s method in two ways: First, data errors are included. It can be easily seen from the proofs, that in (2.20), (2.22) and (2.24), \( b_n \) could be replaced by \( (b_n + \delta_n) \), which yields (formally) weaker conditions. Second, our convergence rates are better than those for Marti’s method. For simplicity, we discuss this aspect for error-free data, i.e., \( \delta_n = 0 \). We first look at part (a) of Theorem 2.6. If we choose \( q = 2 \), we obtain the rate \( O(\sqrt{b_n}) \), which becomes \( O(\sqrt{\gamma_n}) \), if \( b_n \sim \gamma_n \), i.e., if the correct asymptotic behavior of \( \gamma_n \) is known. If we choose higher values of \( q \), we can get arbitrarily close to the optimal rate \( O(\gamma_n) \); e.g., for \( q = 10 \) we can take \( b_n \sim \gamma_n^{5/3} \) and thus obtain the rate \( O(\gamma_n^{5/6}) \).

Of course, if we underestimate \( \gamma_n \), i.e., if we replace \( O \) by \( o \) in (2.22), we get worse rates in terms of \( \gamma_n \), which is not surprising. Note that \( \gamma_n \) represents the best possible convergence rate of elements in \( V_n \) toward solutions of (1.1) that fulfill (1.8). While in Marti’s method, one can obtain at most the square root of the optimal rate (cf. (1.7)), we can come arbitrarily close to the optimal rate, even under the weak smoothness assumption (1.8). In Marti’s method, a better convergence rate cannot be obtained under the stronger assumption (2.23) (cf. the concluding remarks of Section 1); Theorem 2.6(b) shows that we obtain the optimal convergence rate \( O(\gamma_n) \) for \( q \geq 2 \) if we take \( b_n \sim \gamma_n^{1/2} \). In the presence of data errors, the convergence rates are also best possible in terms of \( \delta_n \) (cf., e.g., [1]).

If \( \delta_n = 0 \), one does not need the assumption \( p \leq 2q \), which was needed in Lemma 2.3. In this case, one can show that if (1.8) and (2.20) hold, then \( \|x_n - T^\dagger y\| = O(b_n^{p/2q+2}) \); i.e., if one takes \( b_n \sim \gamma_n^{2q/p} \), then one obtains the rate \( O(\gamma_n^{pAq+2}) \) for any choice of \( p, q > 0 \). In this case, (2.6) reads \( \rho_n(\alpha) = D \cdot (\gamma_n^{2} / \alpha)^q \) with a suitable constant \( D \).

Similarly, if \( \delta_n = 0 \) and (2.24) holds, and if either \( 0 < q < 2 \) and \( \gamma_n = O(b_n^{p/2q}) \) or \( q \geq 2 \) and \( \gamma_n = O(b_n^{Aq+2}) \), then \( \|x_n - T^\dagger y\| = O(b_n^{pAq+2}) \) holds; i.e., if one takes \( b_n \sim \gamma_n^{2q/p} \) for \( q < 2 \) or \( b_n \sim \gamma_n^{A+2/q} \) for \( q \geq 2 \), one obtains the rates \( O(\gamma_n^{2qAq+2}) \) for \( q < 2 \) and the optimal rate \( O(\gamma_n) \) for \( q \geq 2 \). The equation (2.6) then reads \( \rho_n(\alpha) = D \cdot (\gamma_n^{2}/\alpha)^q \) and \( \rho_n(\alpha) = D \cdot \gamma_n^{A+2} \cdot \alpha^{-q} \), respectively (with suitable constants \( D \)).
The proof of these statements follows along the lines of the proof of Theorem 2.6. For actual computations, the choice of $D_1$ and $D_2$ is of course important. Since we want to give convergence rates only, we do not enter into theoretical discussions on the choice of these constants (cf. [3] in a similar situation). In our computations, we chose $D_1 = D_2 = 10^{-6q/p}$, which turned out to be effective.

3. Numerical Aspects and Examples. For computing the solution $x_n^{\delta_n}$ of (2.4) and (2.6), one chooses a basis $\{v_1, \ldots, v_{d(n)}\}$ of $V_n$, computes the $d(n) \times d(n)$-matrices $B_n := (\langle T_{v_i}, T_{v_j} \rangle)$ and $M_n := (\langle v_i, v_j \rangle)$ and the vector $w_n := (\langle T_{v_i}, y_n \rangle)$. It is easy to see that if $(\alpha, \lambda) \in \mathbb{R}^{d(n)+1}$ solves

\begin{align}
(3.1) & \quad (B_n + \alpha M_n) \lambda = w_n, \\
(3.2) & \quad \alpha^{q+2} \lambda^T M_n \lambda = (D_1 b_n + D_2 \delta_n)^p,
\end{align}

then $x_n^{\delta_n} = \sum_{i=1}^{d(n)} \lambda_i v_i$. Note that the system (3.1), (3.2) is very similar to the computational form of Marti’s method (see, e.g., equations (3a), (3b) of [10]). Note that for obtaining (3.2), we used (2.19).

Let for any $\alpha > 0$, $\lambda(\alpha)$ be the unique solution of (3.1), and

\begin{align}
(3.3) & \quad f(\alpha) := \alpha^{q+2} \lambda(\alpha)^T M_n \lambda(\alpha) - (D_1 b_n + D_2 \delta_n)^p. \\
(3.4) & \quad f'(\alpha) = (q + 2) \alpha^{q+1} \lambda(\alpha)^T M_n \lambda(\alpha) \\
& \quad \quad - 2\alpha^{q+2} \lambda(\alpha)^T M_n (B_n + \alpha M_n)^{-1} M_n \lambda(\alpha) > 0;
\end{align}

for all $q \geq 1$, Newton’s method

\begin{align}
(3.5) & \quad \alpha_{k+1} = \alpha_k - \frac{f(\alpha_k)}{f'(\alpha_k)},
\end{align}

converges to the unique zero $\alpha_n$ of $f$ (defined in (3.3)). The convergence is global, for $k \geq 2$ the iterates decrease monotonically to $\alpha_n$. Obviously, the vector $(\alpha_n, \lambda(\alpha_n))$ solves (3.1), (3.2). The solution of (3.1) that is needed in each iteration of (3.5) is done by Cholesky decompositions. For finding a suitable sequence $b_n$, one needs information about $\gamma_n$ (as defined in (2.1)); estimates for $\gamma_n$ for spline spaces can be found, e.g., in [6]. In an analogous way, numerous estimates for the approximating power of finite-element spaces could be used to estimate $\gamma_n$.

In our example, we choose $V_n$ as a space of linear splines on a uniform grid of $(n+1)$ points in $[0,1]$. As basis functions, we take $v_1, \ldots, v_{n+1}$ with the property that $v_i((i-1)/n) = 1$ and $v_i$ vanishes at all other nodes. The elements of $M_n$ are computed explicitly, the elements $T_{v_i}$ are computed by Gaussian quadrature with two nodes on each subinterval $[(i-1)/n, i/n]$. Finally, the scalar products needed for the elements of $B_n$ are approximated by the trapezoidal rule. This is (nearly) identical to the procedure chosen in [9] and [10], so that it is fair to compare the results. All examples are Fredholm integral equations of the first kind on $[0,1]$ with kernel $k$:

\[ \int_0^1 k(t, s) x(s) \, ds = y(s). \]
Example 3.1. Here the kernel is always given by

\[ k(t, s) := \begin{cases} 0, & t \leq s, \\ t - s, & t > s \end{cases} \]

(cf. the example in [10]). It follows from [6] that \( \gamma_n = O(n^{-2}) \).

(a) \( y(t) = \tfrac{1}{3}(6t^5 - 4t^3 + t^2) \). The exact solution is \( (T^*_y)(s) = \tfrac{1}{3}(s - 1)^2 \in R(T^*_y) \), since \( T^*_y = T^*1 \). We use the variant of Theorem 2.6 given in Remark 2.7 under the assumption (1.8) and choose \( p = 1, q = 2, b_n := n^{-2} \). According to the theory, we should obtain the convergence rate \( O(n^{-1}) \). The results are as follows:

<table>
<thead>
<tr>
<th>( n )</th>
<th>( \alpha_n )</th>
<th>( e_n )</th>
<th>( e_n \cdot n \cdot 10^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>( 1.4 \times 10^{-4} )</td>
<td>( 3.4 \times 10^{-3} )</td>
<td>( 1.4 )</td>
</tr>
<tr>
<td>8</td>
<td>( 3.4 \times 10^{-5} )</td>
<td>( 1.1 \times 10^{-4} )</td>
<td>( 0.90 )</td>
</tr>
<tr>
<td>16</td>
<td>( 8.3 \times 10^{-6} )</td>
<td>( 4.3 \times 10^{-5} )</td>
<td>( 0.68 )</td>
</tr>
<tr>
<td>32</td>
<td>( 2.1 \times 10^{-6} )</td>
<td>( 1.7 \times 10^{-5} )</td>
<td>( 0.55 )</td>
</tr>
<tr>
<td>64</td>
<td>( 5.2 \times 10^{-7} )</td>
<td>( 7.1 \times 10^{-6} )</td>
<td>( 0.45 )</td>
</tr>
</tbody>
</table>

Here, \( e_n := \|x_n - T^*_y\| \). The last column shows that the convergence seems to be slightly faster than \( O(n^{-1}) \).

(b) \( y(t) = \tfrac{1}{30}(t^6 - 20t^4 + 45t^2) \), \( (T^*_y)(s) = \tfrac{1}{30}(s^4 - 4s^3 + 3) \in R(T^*_T) \), since \( T^*_y = T^*T_1 \). Choosing the parameters as in (a), the results outlined in Remark 2.7 show that the convergence rate should be the optimal rate \( O(n^{-2}) \). This is in fact the case:

<table>
<thead>
<tr>
<th>( n )</th>
<th>( \alpha_n )</th>
<th>( e_n )</th>
<th>( e_n \cdot n^2 \cdot 10^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>( 2.5 \times 10^{-4} )</td>
<td>( 9.1 \times 10^{-4} )</td>
<td>( 1.5 )</td>
</tr>
<tr>
<td>8</td>
<td>( 6.3 \times 10^{-5} )</td>
<td>( 2.2 \times 10^{-4} )</td>
<td>( 1.4 )</td>
</tr>
<tr>
<td>16</td>
<td>( 1.6 \times 10^{-5} )</td>
<td>( 5.2 \times 10^{-5} )</td>
<td>( 1.3 )</td>
</tr>
<tr>
<td>32</td>
<td>( 3.9 \times 10^{-6} )</td>
<td>( 1.2 \times 10^{-5} )</td>
<td>( 1.2 )</td>
</tr>
<tr>
<td>64</td>
<td>( 9.7 \times 10^{-7} )</td>
<td>( 2.9 \times 10^{-6} )</td>
<td>( 1.2 )</td>
</tr>
</tbody>
</table>

(c) \( y \) as in (a); \( y \) is randomly perturbed by \( \delta_n = n^{-2}\|y\| \). We use Theorem 2.6(a) with \( b_n = n^{-2} \), \( p = 4, q = 2 \). According to the theory, we should obtain the rate \( \|x_n^{\delta_n} - T^*_y\| = O(n^{-1}) = O(\delta_n^{1/2}) \).

<table>
<thead>
<tr>
<th>( n )</th>
<th>( \alpha_n )</th>
<th>( \delta_n )</th>
<th>( \delta_n \cdot n \cdot 10 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>( 1.4 \times 10^{-4} )</td>
<td>( 5.6 \times 10^{-2} )</td>
<td>( 2.2 )</td>
</tr>
<tr>
<td>8</td>
<td>( 3.4 \times 10^{-5} )</td>
<td>( 7.9 \times 10^{-2} )</td>
<td>( 6.3 )</td>
</tr>
<tr>
<td>16</td>
<td>( 8.8 \times 10^{-6} )</td>
<td>( 3.1 \times 10^{-2} )</td>
<td>( 4.9 )</td>
</tr>
<tr>
<td>32</td>
<td>( 2.2 \times 10^{-6} )</td>
<td>( 1.3 \times 10^{-2} )</td>
<td>( 4.2 )</td>
</tr>
<tr>
<td>64</td>
<td>( 5.5 \times 10^{-7} )</td>
<td>( 5.9 \times 10^{-3} )</td>
<td>( 3.8 )</td>
</tr>
</tbody>
</table>

Here (and below), \( \delta_n := \|x_n^{\delta_n} - T^*_y\| \).
(d) $y$ as in (b); $y$ is randomly perturbed by $\delta_n = n^{-3}||y||$. We use Theorem 2.6(b) with $b_n = n^{-3}$, $p = 8/3$, $q = 2$. According to the theory, we should obtain $||x_n^d - T^*y|| = O(n^{-2}) = O(\delta_n^{2/3})$.

<table>
<thead>
<tr>
<th>$n$</th>
<th>$\alpha_n$</th>
<th>$\tilde{e}_n$</th>
<th>$\tilde{e}_n \cdot n^2 \cdot 10^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>$2.5 \cdot 10^{-4}$</td>
<td>$3.6 \cdot 10^{-3}$</td>
<td>5.8</td>
</tr>
<tr>
<td>8</td>
<td>$6.1 \cdot 10^{-5}$</td>
<td>$2.1 \cdot 10^{-3}$</td>
<td>13</td>
</tr>
<tr>
<td>16</td>
<td>$1.5 \cdot 10^{-5}$</td>
<td>$3.6 \cdot 10^{-4}$</td>
<td>9.3</td>
</tr>
<tr>
<td>32</td>
<td>$3.8 \cdot 10^{-6}$</td>
<td>$8.4 \cdot 10^{-5}$</td>
<td>8.6</td>
</tr>
<tr>
<td>64</td>
<td>$9.5 \cdot 10^{-7}$</td>
<td>$2.1 \cdot 10^{-5}$</td>
<td>8.6</td>
</tr>
</tbody>
</table>

Computations with larger values for $\delta_n$ give comparable results.

**Example 3.2.** Here the kernel is given by

$$k(t, s) := \begin{cases} t(1 - s), & t \leq s, \\ s(1 - t), & t > s \end{cases}$$

(cf. Example b) in [9]). Again, $\gamma_n = O(n^{-2})$.

In parts (a)–(d) of this example, the same choices of the relevant parameters are made and the same statements about convergence rates hold as in the corresponding parts of Example 3.1.

(a) $y(t) = \frac{1}{15}(t - 2t^3 + t^4), (T^*y)(s) = \frac{1}{2}(s - s^2)$.

<table>
<thead>
<tr>
<th>$n$</th>
<th>$\alpha_n$</th>
<th>$e_n$</th>
<th>$e_n \cdot n \cdot 10^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>$2.1 \cdot 10^{-4}$</td>
<td>$4.0 \cdot 10^{-3}$</td>
<td>1.6</td>
</tr>
<tr>
<td>8</td>
<td>$5.2 \cdot 10^{-5}$</td>
<td>$1.4 \cdot 10^{-3}$</td>
<td>1.1</td>
</tr>
<tr>
<td>16</td>
<td>$1.3 \cdot 10^{-5}$</td>
<td>$5.5 \cdot 10^{-4}$</td>
<td>0.88</td>
</tr>
<tr>
<td>32</td>
<td>$3.3 \cdot 10^{-6}$</td>
<td>$2.2 \cdot 10^{-4}$</td>
<td>0.72</td>
</tr>
<tr>
<td>64</td>
<td>$8.1 \cdot 10^{-7}$</td>
<td>$9.3 \cdot 10^{-5}$</td>
<td>0.60</td>
</tr>
</tbody>
</table>

(b) $y(t) = \frac{1}{30}(3t - 5t^3 + 3t^5 + t^6), (T^*y)(s) = s - 2s^3 + s^4$.

<table>
<thead>
<tr>
<th>$n$</th>
<th>$\alpha_n$</th>
<th>$e_n$</th>
<th>$e_n \cdot n^2 \cdot 10^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>$1.4 \cdot 10^{-4}$</td>
<td>$6.2 \cdot 10^{-3}$</td>
<td>10.0</td>
</tr>
<tr>
<td>8</td>
<td>$3.4 \cdot 10^{-5}$</td>
<td>$1.5 \cdot 10^{-3}$</td>
<td>9.6</td>
</tr>
<tr>
<td>16</td>
<td>$8.3 \cdot 10^{-6}$</td>
<td>$3.7 \cdot 10^{-4}$</td>
<td>9.5</td>
</tr>
<tr>
<td>32</td>
<td>$2.1 \cdot 10^{-6}$</td>
<td>$9.3 \cdot 10^{-5}$</td>
<td>9.5</td>
</tr>
<tr>
<td>64</td>
<td>$5.2 \cdot 10^{-7}$</td>
<td>$2.3 \cdot 10^{-5}$</td>
<td>9.6</td>
</tr>
</tbody>
</table>
ILL-POSED LINEAR INTEGRAL EQUATIONS 415

(c) $y$ as in (a); $y$ is randomly perturbed by $\delta_n = n^{-2}\|y\|$. 

<table>
<thead>
<tr>
<th>$n$</th>
<th>$\alpha_n$</th>
<th>$\hat{e}_n$</th>
<th>$\hat{e}_n \cdot n \cdot 10^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>$2.0 \times 10^{-4}$</td>
<td>$8.3 \times 10^{-3}$</td>
<td>3.3</td>
</tr>
<tr>
<td>8</td>
<td>$5.3 \times 10^{-5}$</td>
<td>$1.5 \times 10^{-2}$</td>
<td>12.4</td>
</tr>
<tr>
<td>16</td>
<td>$1.3 \times 10^{-5}$</td>
<td>$6.2 \times 10^{-3}$</td>
<td>9.8</td>
</tr>
<tr>
<td>32</td>
<td>$3.3 \times 10^{-6}$</td>
<td>$2.6 \times 10^{-3}$</td>
<td>8.2</td>
</tr>
<tr>
<td>64</td>
<td>$8.3 \times 10^{-7}$</td>
<td>$1.3 \times 10^{-3}$</td>
<td>8.3</td>
</tr>
</tbody>
</table>

(d) $y$ as in (b), $y$ is randomly perturbed by $\delta_n = n^{-3}\|y\|$. 

<table>
<thead>
<tr>
<th>$n$</th>
<th>$\alpha_n$</th>
<th>$\hat{e}_n$</th>
<th>$\hat{e}_n \cdot n^2 \cdot 10$</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>$1.3 \times 10^{-4}$</td>
<td>$5.8 \times 10^{-3}$</td>
<td>0.93</td>
</tr>
<tr>
<td>8</td>
<td>$3.3 \times 10^{-5}$</td>
<td>$3.8 \times 10^{-3}$</td>
<td>2.4</td>
</tr>
<tr>
<td>16</td>
<td>$8.2 \times 10^{-6}$</td>
<td>$8.3 \times 10^{-4}$</td>
<td>2.1</td>
</tr>
<tr>
<td>32</td>
<td>$2.0 \times 10^{-6}$</td>
<td>$1.8 \times 10^{-4}$</td>
<td>1.8</td>
</tr>
<tr>
<td>64</td>
<td>$5.1 \times 10^{-7}$</td>
<td>$4.1 \times 10^{-5}$</td>
<td>1.7</td>
</tr>
</tbody>
</table>

Again, computations with larger values for $\delta_n$ give comparable results.

(e) $y(t) = \frac{1}{t} (t - t^3), (T^* y)(s) = s$. Note that (1.8) is not fulfilled. Nevertheless, we use the same parameters (except for the values of $D_1$ and $D_2$) as in (a) and obtain:

<table>
<thead>
<tr>
<th>$n$</th>
<th>$\alpha_n$</th>
<th>$e_n$</th>
<th>$e_n \cdot n^{1/2} \cdot 10$</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>$2.8 \times 10^{-6}$</td>
<td>$2.7 \times 10^{-1}$</td>
<td>5.37</td>
</tr>
<tr>
<td>8</td>
<td>$6.8 \times 10^{-7}$</td>
<td>$1.9 \times 10^{-1}$</td>
<td>5.38</td>
</tr>
<tr>
<td>16</td>
<td>$1.7 \times 10^{-7}$</td>
<td>$1.4 \times 10^{-1}$</td>
<td>5.40</td>
</tr>
<tr>
<td>32</td>
<td>$4.1 \times 10^{-8}$</td>
<td>$9.9 \times 10^{-2}$</td>
<td>5.58</td>
</tr>
</tbody>
</table>

The last column shows that instead of the rate $O(n^{-1})$, which would be achieved if (1.8) were fulfilled, the rate seems to be $O(n^{-1/2})$ here. Incidentally, the errors are significantly larger than in (a) and a bit (but not much) smaller than in [9, Example c]. This confirms Marti’s observation that for this example, Tikhonov regularization using the $L^2$-norm is not good. If one uses Tikhonov regularization using $\|x''\|^2$ instead of $\|x\|^2$, the results are better (cf. [3, Example 2]).

Institut für Mathematik
Johannes-Kepler-Universität
A-4040 Linz, Austria