MATHEMATICS OF COMPUTATION
VOLUME 49, NUMBER 180
OCTOBER 1987, PAGES S5–S12

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

By Helmut Gfrerer

APPENDIX 1. SINGULAR SYSTEMS

Some of the proofs in Appendix 2 are based on the use of a singular system $\{\sigma_j, u_j, v_j\}$ for the compact operator T (cf. [8], [12]). We now give a brief summary of the most important preliminaries used in Appendix 2.

From spectral theory we know that the nonzero eigenvalues of T^*T can be enumerated as a sequence $\sigma_1^2 \geq \sigma_2^2 \geq \ldots$ which (if infinite) converges to zero. If we denote by $u_1, u_2, \ldots$ an associated sequence of orthonormal eigenvectors and set $v_j = Tu_j/\sigma_j$, then $T^*v_j = \sigma_j u_j$. Moreover, $\{u_n\}$ is a complete orthonormal set for $\overline{R(T^*)} = N(T)^\perp$ and $\{v_n\}$ is a complete orthonormal set for $\overline{R(T)} = N(T^*)^\perp$. In order that $y \in D(T^\dagger)$, it is necessary and sufficient that $\sum_j \sigma_j^{-2}\langle y, v_j\rangle^2 < \infty$. Then $T^\dagger y = \sum_j \sigma_j^{-1}\langle y, v_j\rangle u_j$. Further, $T^\dagger y \in R((T^*T)^\nu)$ if and only if $\sum_j \sigma_j^{-(2+4\nu)}\langle y, v_j\rangle^2 < \infty$.
For any $\alpha, \lambda > 0$ we have for each $z \in Y$

$$(\alpha I + TT^*)^{-\lambda} z = \sum_j (\alpha + \sigma_j^2)^{-\lambda}\langle z, v_j\rangle v_j + \alpha^{-\lambda}(I-Q)z,$$

where Q denotes the orthogonal projector onto $\overline{R(T)}$. Hence

$$\|\alpha^\lambda(\alpha I + TT^*)^{-\lambda} z\|^2 = \sum_j \alpha^{2\lambda}(\alpha+\sigma_j^2)^{-2\lambda}\langle z, v_j\rangle^2 + \|(I-Q)z\|^2 \leq \sum_j \langle z, v_j\rangle^2 + \|(I-Q)z\|^2 = \|z\|^2,$$

and this implies $\|\alpha^\lambda(\alpha I + TT^*)^{-\lambda}\| \leq 1$.
Analogously, we have for each $x \in X$

$$(\alpha I + T^*T)^{-\lambda} x = \sum_j (\alpha + \sigma_j^2)^{-\lambda}\langle x, u_j\rangle u_j + \alpha^{-\lambda}(I-P)x,$$

where P is the orthogonal projector onto $\overline{R(T^*)}$. Hence, our approximations given by (1.3) may be written as

$$x_{\alpha,\delta} = (\alpha I+T^*T)^{-1}T^*y_\delta = \sum_j (\alpha+\sigma_j^2)^{-1}\sigma_j\langle y_\delta,v_j\rangle u_j.$$

We now formulate a lemma which will be useful in the sequel.

Lemma A.1. *Let* $\sum_{j=1}^{\infty} b_j^2 < \infty$ *and* $\{g_j(\alpha)\}$ *be a sequence of continuous functions on* $[0,\infty)$, *uniformly bounded in* j *and* α. *Further, let* $g_j(0) = 0$ *for each* j. *Then the series* $\sum_{j=1}^{\infty} g_jb_j^2$ *converges uniformly to a continuous function* G *on* $[0,\infty)$ *with* $\lim_{\alpha\to 0} G(\alpha) = G(0) = 0$.

The proof is straightforward and is omitted.

APPENDIX 2. PROOFS.

Proof of Lemma 2.1. If $\{\sigma_j,u_j,v_j\}$ denotes a singular system, we obtain

$$f_n(\alpha,z) = \sum_j \alpha^{2n+1}(\alpha+\sigma_j^2)^{-(2n+1)}\langle z,v_j\rangle^2.$$

Then, by Lemma A.1, $f_n(\alpha,z)$ is continuous and $\lim_{\alpha\to 0} f_n(\alpha,z) = 0$. Since $\alpha^{2n+1}(\alpha+\sigma_j^2)^{-(2n+1)}$ is strictly increasing for each j and $\|Qz\|^2 = \sum_j\langle z,v_j\rangle^2 \neq 0$, we obtain that $f_n(\alpha,z)$ is strictly increasing.

Since the sequence $\{\sigma_j^2\}$ is decreasing, we have

$$0 \le \|Qz\|^2-f_n(\alpha,z) = \sum_j (1-\alpha^{2n+1}(\alpha+\sigma_j^2)^{-(2n+1)})\langle z,v_j\rangle^2$$

$$\le (1-\alpha^{2n+1}(\alpha+\sigma_1^2)^{-(2n+1)}) \sum_j \langle z,v_j^2\rangle \to 0 \quad \text{for } \alpha \to \infty.$$

Thus, $\lim_{\alpha\to\infty} f_n(\alpha,z) = \|Qz\|^2$.

It follows immediately from (1.7) that

$$x_\alpha^n = \sum_{k=1}^{n} \alpha^{k-1}(\alpha I+T^*T)^{-k}T^*y = \sum_{k=1}^{n}\sum_j \alpha^{k-1}(\alpha+\sigma_j^2)^{-k}\sigma_j\langle y,v_j\rangle u_j$$

$$= \sum_j (1-\alpha^n(\alpha+\sigma_j^2)^{-n})\sigma_j^{-1}\langle y,v_j\rangle u_j.$$

Hence, we obtain $T^\dagger y-x_\alpha^n = \sum_j \alpha^n(\alpha+\sigma_j^2)^{-n}\sigma_j^{-1}\langle y,v_j\rangle u_j$ and

$$\varphi_n(\alpha) = \sum_j \alpha^{2n}(\alpha+\sigma_j^2)^{-2n}\sigma_j^{-2}\langle y,v_j\rangle^2.$$

It follows, analogously as above, that φ_n is strictly increasing, continuous and $\lim_{\alpha\to 0}\varphi_n(\alpha) = 0$, $\lim_{\alpha\to\infty}\varphi_n(\alpha) = \sum_j \sigma_j^{-2}\langle y,v_j\rangle^2 = \|T^\dagger y\|^2$. Further, we obtain by standard arguments that φ_n is continuously differentiable and

$$\varphi_n'(\alpha) = 2n\sum_j \alpha^{2n-1}(\alpha+\sigma_j^2)^{-(2n+1)}\langle y,v_j\rangle^2 = 2n\alpha^{-2} f_n(\alpha,y).$$

□

<u>Proof of Lemma 2.5.</u> For $\alpha > 0$ we define the linear operator $\phi_\alpha^n := \alpha^{(2n+1)/2}(\alpha I+TT^*)^{-(2n+1)/2}Q$. Since $\|\phi_\alpha^n\| \le |\alpha^{(2n+1)/2}(\alpha I+TT^*)^{-(2n+1)/2}\|\|Q\| \le 1$, we obtain

$$\|\phi^n_{\alpha(\delta)}y\| \le \|\phi^n_{\alpha(\delta)}y_\delta\| + \|\phi^n_{\alpha(\delta)}(y-y_\delta)\| \le C^{1/2}\,\delta + \delta$$

and

$$\|\phi^n_{\alpha(\delta)}y\| \ge \|\phi^n_{\alpha(\delta)}y_\delta\| - \|\phi^n_{\alpha(\delta)}(y-y_\delta)\| \ge C^{1/2}\,\delta - \delta.$$

Since $f_n(\alpha(\delta),y) = \|\phi^n_{\alpha(\delta)}y\|^2$, our assertion follows immediately. □

<u>Proof of Lemma 2.6.</u> Since $|t \sum_{k=1}^{n} \alpha^{k-1}(\alpha+t)^{-k}| = |1-\alpha^n(\alpha+t)^{-n}| \le 1$ and $|\sum_{k=1}^{n} \alpha^{k-1}(\alpha+t)^{-k}| \le n/\alpha$ hold for all $\alpha > 0$, $t \ge 0$, we obtain from [12, Lemma 2.3.2] that

$$\|x^n_{\alpha,\delta}-T^\dagger y\| \le \|x^n_\alpha-T^\dagger y\| + \delta(n/\alpha)^{1/2}.$$

There follows

$$\varphi_{n,\delta}(\alpha) \le (\|x^n_\alpha-T^\dagger y\| + \delta(n/\alpha)^{1/2})^2 \le 2\,(2n\delta^2/\alpha + \varphi_n(\alpha))$$

$$\le 2/\min\{\gamma,1\}\;(2n\gamma\delta^2/\alpha + \varphi_n(\alpha)).$$ □

<u>Proof of Lemma 2.7.</u> Denote $g(\alpha) := 2n\gamma\delta^2/\alpha + \varphi_n(\alpha)$. A minimizer $\bar\alpha$ of (2.7) has to satisfy the first-order necessary condition $\frac{d}{d\alpha}\,g(\bar\alpha) = 0$. Using Lemma 2.1, this is equivalent to $(-\gamma\delta^2 + f_n(\bar\alpha,y))2n/\bar\alpha^2 = 0$.

Now suppose that $f_n(\bar\alpha,y) = \gamma\delta^2$ for some $\bar\alpha > 0$. Using the monotonicity of $f_n(\cdot,y)$ we have $\frac{d}{d\alpha}\,g(\alpha) = (-\gamma\,^2+f_n(\alpha,y))2n/\alpha^2 > 0$ for $\alpha > \bar\alpha$ and $\frac{d}{d\alpha}\,g(\alpha) < 0$ for $0 < \alpha < \bar\alpha$. Thus $\bar\alpha$ is a minimizer for (2.7). □

<u>Proof of Theorem 2.9.</u> The first part follows immediately from (2.1o) and (2.9). To prove the second part, we will show that $\varphi_n(n\alpha) \le \varphi_m(m\alpha)$ for $m < n$ and all $\alpha > 0$.

Using a singular system$\{\sigma_j;\, u_j;\, v_j\}$ for T, we have for all $n \in \mathbf{N}$

(A.1) $$\varphi_n(\alpha) = \sum_j \alpha^{2n}(\alpha+\sigma_j^2)^{-2n}\,\sigma_j^{-2}\langle y,v_j\rangle^2.$$

Let $g_j(\nu) := [\nu\alpha/(\nu\alpha+\sigma_j^2)]^{2\nu}$ for $\nu > 0$. Since

$g_j'(\nu) = 2[\nu\alpha/(\nu\alpha+\sigma_j^2)]^{2\nu}[\sigma_j^2/(\nu\alpha+\sigma_j^2) + \ln(\nu\alpha/(\nu\alpha+\sigma_j^2))]$ and $1-x + \ln x \le 0$ for all $x > 0$, we obtain $g_j'(\nu) \le 0$. Thus, g_j is decreasing and hence $g_j(n) \le g_j(m)$ for $m < n$. Thus it follows from (A.1) that

$$\varphi_n(n\alpha) = \sum_j g_j(n)\,\sigma_j^{-2}\langle y,v_j\rangle^2 \le \sum_j g_j(m)\sigma_j^{-2}\langle y,v_j\rangle^2 = \varphi_m(m\alpha).$$

Together with (2.8), there follows

$$2m\delta^2 = \varphi_m(\beta_m(\delta))\beta_m(\delta) = \varphi_n(\beta_n(\delta))\beta_n(\delta)m/n \le \varphi_m(\beta_n(\delta)m/n)\beta_n(\delta)m/n.$$

Since the function $\beta\,\varphi_m(\beta)$ is increasing, we obtain $\beta_m(\delta) \le \beta_n(\delta)\,m/n$ for $m < n$.

Hence, $\varphi_m(\beta_m(\delta)) = 2m\,\delta^2/\beta_m(\delta) \ge 2n\,\delta^2/\beta_n(\delta) = \varphi_n(\beta_n(\delta))$. □

<u>Proof of Theorem 2.1o.</u> Let $\{\sigma_j;\, u_j;\, v_j\}$ be a singular system for T and define

(A.2) $$\psi(\alpha) := \sum_j \alpha(\alpha+\sigma_j^2)^{-1}\,\sigma_j^{-2}\langle y,v_j\rangle^2.$$

Since $\varphi_1(\alpha) = \sum_j \alpha^2(\alpha+\sigma_j^2)^{-2}\,\sigma_j^{-2}\langle y,v_j\rangle^2$, we have $\varphi_1(\alpha) \le \psi(\alpha)$. Further,

$$\text{(A.3)}\qquad \|(\alpha I+TT^*)^{-1}Qy\|^2 = \sum_j (\alpha+\sigma_j^2)^{-2}\langle y,v_j\rangle^2 \le \alpha^{-1}\sum_j \alpha(\alpha+\sigma_j^2)^{-1}\sigma_j^{-2}\langle y,v_j\rangle^2 = \alpha^{-1}\psi(\alpha).$$

Because of $\delta\alpha^{-1/2}\psi(\alpha)^{1/2} \le (\delta^2\alpha^{-1}+\psi(\alpha))/2$ we obtain by (2.14) and Lemma 2.6

$$\text{(A.4)}\qquad \|x_{\alpha,\delta}-T^\dagger y\|^2 \le E(\alpha,y_\delta) \le 2(2\delta^2/\alpha+\varphi_1(\alpha)) + 4\delta\alpha^{-1/2}\psi(\alpha)^{1/2} + 8\delta^2/\alpha \le 4\psi(\alpha) + 14\delta^2/\alpha.$$

Now suppose that $T^\dagger y \in R((T^*T)^\nu)$. Since this holds if and only if $\sum_j \sigma_j^{-(2+4\nu)}\langle y,v_j\rangle^2 < \infty$, we obtain

$$\text{(A.5)}\qquad \psi(\alpha) \le \sum_j \alpha\sigma_j^{-4}\langle y,v_j\rangle^2 = O(\alpha) \quad \text{for } \nu \ge 1/2,$$

and by Lemma A.1,

$$\text{(A.6)}\qquad \psi(\alpha) \le \alpha^{2\nu}\sum_j \alpha^{1-2\nu}(\alpha+\sigma_j^2)^{-(1-2\nu)}(\alpha+\sigma_j^2)^{-2\nu}\sigma_j^{-2}\langle y,v_j\rangle^2$$
$$\le \alpha^{2\nu}\sum_j \alpha^{1-2\nu}(\alpha+\sigma_j^2)^{-(1-2\nu)}\sigma_j^{-(2+4\nu)}\langle y,v_j\rangle^2 = o(\alpha^{2\nu}) \text{ for } \nu < 1/2.$$

It can easily be shown that for each $\delta > 0$ the equation

$$\psi(\beta) = \delta^2/\beta$$

has a unique solution $\beta = \beta(\delta)$.

Since $\alpha(\delta)$ minimizes $E(\alpha,y_\delta)$, we obtain from (A.4)

$$\|x_{\alpha(\delta),\delta}-T^\dagger y\|^2 \le E(\alpha(\delta),y_\delta) = \min\{E(\alpha,y_\delta): \alpha > 0\} \le \inf\{4\psi(\alpha)+14\delta^2/\alpha: \alpha > 0\} \le 4\psi(\beta(\delta)) + 14\delta^2/\beta(\delta).$$

Using (A.5) and (A.6), one can show, analogously to the proof of Theorem 2.8, that $\delta^2/\beta(\delta) = \psi(\beta(\delta)) = O(\delta)$ for $\nu \ge 1/2$ and $\delta^2/\beta(\delta) = \psi(\beta(\delta)) = o(\delta^{4\nu/(2\nu+1)})$ for $\nu < 1/2$. □

<u>Proof of Lemma 3.2.</u> Analogously to the proof of Lemma 2.1, one shows that $f_n^m(\alpha,y_\delta)$ is continuous, strictly increasing, and $\lim_{\alpha\to 0} f_n^m(\alpha,y_\delta) = 0$, $\lim_{\alpha\to\infty} f_n^m(\alpha,y_\delta) = \|Q_m y_\delta\|^2$. Further, $(1-K(2n-1)b_m^2/\alpha)$ is continuous, increasing and positive for $\alpha > \bar{\alpha} := K(2n-1)b_m^2$. Hence, $(1-K(2n-1)b_m^2/\alpha)\, f_n^m(\alpha,y_\delta)$ is continuous, strictly increasing for $\alpha > \bar{\alpha}$ and $\lim_{\alpha\to\bar{\alpha}} (1-K(2n-1)b_m^2/\alpha)\, f_n^m(\alpha,y_\delta) = 0$, $\lim_{\alpha\to\infty} (1-K(2n-1)b_m^2/\alpha)\, f_n^m(\alpha,y_\delta) = \|Q_m y_\delta\|^2$. Thus, the assertion follows from the Intermediate Value Theorem, resp., if $\delta = 0$, from the fact, that $\bar{\alpha} > 0$. □

<u>Proof of Lemma 3.4.</u> Since the operators $(\alpha I+T_m^*T_m)^{-1}$ and P_m commute, and $T_m^* = P_m T^*$, we have for each $j > 1$

$$\text{(A.7)}\qquad z_{\alpha,m}^j - z_\alpha^j = z_{\alpha,m}^j - \alpha(\alpha I+T_m^*T_m)^{-1}P_m\, z_\alpha^{j-1} + \alpha(\alpha I+T_m^*T_m)^{-1}P_m\, z_\alpha^{j-1} - z_\alpha^j$$
$$= (\alpha I+T_m^*T_m)^{-1}P_m(z_{\alpha,m}^{j-1}-z_\alpha^{j-1}) + (\alpha I+T_m^*T_m)^{-1}[P_m(\alpha I+T^*T) - (\alpha I+T_m^*T_m)]\, z_\alpha^j.$$

Further, we get

(A.8) $$P_m(\alpha I+T^*T)-(\alpha I+T_m^*T_m) = \alpha(P_m-I)+T_m^*T-T_m^*T_m = (T_m^*T-\alpha I)(I-P_m)$$

$$= (T_m^*Q_mT-\alpha I)(I-P_m).$$

Since $\|\alpha(\alpha I+T_m^*T_m)^{-1}\| \le 1$ and $\|(\alpha I+T_m^*T_m)^{-1}\, T_m^*\| \le \alpha^{-1/2}/2$, we obtain from (A.7) and (A.8), for each $j > 1$,

(A.9) $$\|z_{\alpha,m}^j-z_\alpha^j\| \le \|z_{\alpha,m}^{j-1}-z_\alpha^{j-1}\| + \|\alpha(\alpha I+T_m^*T_m)^{-1}(I-P_m)z_\alpha^j\|$$

$$+ \|(\alpha I+T_m^*T_m)^{-1}T_m^*Q_m\, T(I-P_m)z_\alpha^j\| \le \|z_{\alpha,m}^{j-1}-z_\alpha^{j-1}\| + (1+b_m\alpha^{-1/2}/2)\, \|(I-P_m)z_\alpha^j\|.$$

Since $\|z_{\alpha,m}^1-z_\alpha^1\| = \|(\alpha I+T_m^*T_m)^{-1}[P_m(\alpha I+T^*T)-(\alpha I+T_m^*T_m)]z_\alpha^1\| \le (1+b_m\alpha^{-1/2}/2)\|(I-P_m)z_\alpha^1\|$, we obtain by (A.9), for each $j \in \mathbb{N}$,

(A.10) $$\|z_{\alpha,m}^j-z_\alpha^j\| \le \sum_{i=1}^{j} (1+b_m\alpha^{-1/2}/2)\|(I-P_m)z_\alpha^i\|.$$

Because of $x_{\alpha,m}^n-x_\alpha^n = \sum_{j=1}^{n}(z_{\alpha,m}^j-z_\alpha^j)$, this implies

$$\|x_{\alpha,m}^n-x_\alpha^n\| \le \sum_{j=1}^{n}\|z_{\alpha,m}^j-z_\alpha^j\| \le \sum_{j=1}^{n}\sum_{i=1}^{j} (1+b_m\alpha^{-1/2}/2)\|(I-P_m)z_\alpha^i\|$$

$$= \sum_{j=1}^{n}(n-j+1)(1+b_m\alpha^{-1/2}/2)\|(I-P_m)z_\alpha^j\|.$$

□

<u>Proof of Lemma 3.5.</u> Since $\|\alpha(\alpha I+TT^*)^{-1}\| \le 1$ and $\alpha \ge \gamma_m^2$, we obtain, for each $j \in \mathbb{N}$,

$$\|(I-P_m)z_\alpha^j\| = \|(I-P_m)T^*\alpha^{j-1}(\alpha I+TT^*)^{-j}\, Qy\| \le \gamma_m\|(\alpha I+TT^*)^{-1}Qy\| \le \gamma_m\|(\gamma_m^2I+TT^*)^{-1}Qy\|.$$

This, together with (A.2) and (A.3), implies $\|(I-P_m)z_\alpha^j\|^2 \le \psi(\gamma_m^2)$.

Since $b_m \le \gamma_m$, the assertions follow from (A.5), (A.6) and Lemma 3.4. □

<u>Proof of Lemma 3.7.</u> A simple calculation gives

(A.11) $$\|x_{\alpha,m}^n-T^\dagger y\|^2 = \|x_{\alpha,m}^n-T_m^\dagger y\|^2 + 2\langle x_{\alpha,m}^n,T_m^\dagger y-T^\dagger y\rangle - 2\langle T_m^\dagger y,T_m^\dagger y-T^\dagger y\rangle$$

$$+ \|T_m^\dagger y-T^\dagger y\|^2.$$

Since $T_m^*Q_m = T_m^*$ and $Q_m = Q_mQ$, we obtain

$$2\,|\langle x_{\alpha,m}^n,T_m^\dagger y-T^\dagger y\rangle| = 2\,|\langle \sum_{j=1}^{n}\alpha^{j-1}(\alpha I+T_m^*T_m)^{-j}T_m^*y,T_m^\dagger y-T^\dagger y\rangle|$$

$$= 2\,|\langle \sum_{j=1}^{n}\alpha^{j-1}(\alpha I+T_mT_m^*)^{-j}Q_my,Q_my-T_mT^\dagger y\rangle|$$

$$= 2\,|\langle \sum_{j=1}^{n}\alpha^{j-1}(\alpha I+T_mT_m^*)^{-j}Q_my,Q_m(T-T_m)T^\dagger y\rangle|$$

$$\le 2b_m\|\sum_{j=1}^{n}\alpha^{j-1}(\alpha I+T_mT_m^*)^{-j}Q_my\|\,\|(I-P_m)T^\dagger y\|$$

$$\le Kb_m^2\|\sum_{j=1}^{n}\alpha^{j-1}(\alpha I+T_mT_m^*)^{-j}Q_my\|^2 + (1/K)\|(I-P_m)T^\dagger y\|^2.$$

One can easily prove by induction that

$$\Big\| \sum_{j=1}^{n} \alpha^{j-1}(\alpha I+T_mT_m^*)^{-j}Q_my\Big\|^2 = \sum_{j=1}^{n} j\rho_{j,m}(\alpha) + \sum_{j=n+1}^{2n-1} (2n-j)\rho_{j,m}(\alpha).$$

Since $\rho_{j,m}(\alpha) = \alpha^{j-1}\|(\alpha I+T_mT_m^*)^{-(j+1)/2}Q_my\|^2 \geq 0$, we obtain

$$\Big\| \sum_{j=1}^{n} \alpha^{j-1}(\alpha I+T_mT_m^*)^{-j}Q_my\Big\|^2 \leq \sum_{j=1}^{2n-1} j\rho_{j,m}(\alpha).$$

This, together with (A.12), yields

$$\text{(A.13)}\qquad 2\,|\langle x^n_{\alpha,m},T_m^{\dagger}y-T^{\dagger}y\rangle| \leq Kb_m^2 \sum_{j=1}^{2n-1} j\rho_{j,m}(\alpha) + (1/K)\|(I-P_m)T^{\dagger}y\|^2.$$

Together with (A.11), this implies

$$\text{(A.14)}\qquad \|x^n_{\alpha,m}-T^{\dagger}y\|^2 \leq \|x^n_{\alpha,m}-T_m^{\dagger}y\|^2 - 2\langle T_m^{\dagger}y,T_m^{\dagger}y-T^{\dagger}y\rangle + \|T_m^{\dagger}y-T^{\dagger}y\|^2$$

$$+ Kb_m^2 \sum_{j=1}^{2n-1} j\rho_{j,m}(\alpha) + (1/K)\|(I-P_m)T^{\dagger}y\|^2.$$

Further, (A.11) implies

$$\|x^n_{\alpha,m}-T_m^{\dagger}y\|^2 - 2\langle T_m^{\dagger}y,T_m^{\dagger}y-T^{\dagger}y\rangle + \|T_m^{\dagger}y-T^{\dagger}y\|^2 \leq \|x^n_{\alpha,m}-T^{\dagger}y\|^2 + 2|\langle x^n_{\alpha,m},T_m^{\dagger}y-T^{\dagger}y\rangle|,$$

and this, together with (A.13) and (A.14), yields the second part of the inequality. □

Proof of Lemma 3.8. Since $\|\alpha^{(j-1)/2}(\alpha I+T_mT_m^*)^{-(j-1)/2}\| \leq 1$, we obtain

$$\rho_{j,m}(\alpha) = \|\alpha^{(j-1)/2}(\alpha I+T_mT_m^*)^{-(j-1)/2}(\alpha I+T_mT_m^*)^{-1}Q_my\|^2 \leq \|(\alpha I+T_mT_m^*)^{-1}Q_my\|^2.$$

Since $Q_m = Q_mQ$ and $\|TT^*-T_mT_m^*\| = \|T(I-P_m)T^*\| = \gamma_m^2$, we obtain

$$\|(\alpha I+T_mT_m^*)^{-1}Q_my\| \leq \|(\alpha I+T_mT_m^*)^{-1}Q_my-(\alpha I+TT^*)^{-1}Qy\| + \|(\alpha I+TT^*)^{-1}Qy\|$$

$$= \|(\alpha I+T_mT_m^*)^{-1}[\alpha(Q_m-I)+Q_m(TT^*-T_mT_m^*)](\alpha I+TT^*)^{-1}Qy\| + \|(\alpha I+TT^*)^{-1}Qy\|$$

$$\leq (1+\gamma_m^2/\alpha)\,\|(\alpha I+TT^*)^{-1}Qy\| + \|(\alpha I+TT^*)^{-1}Qy\| = (2+\gamma_m^2/\alpha)\|(\alpha I+TT^*)^{-1}Qy\|.$$

Thus, $b_m^2\rho_{j,m}(\alpha) \leq b_m^2(2+\gamma_m^2/\alpha)^2\|(\alpha I+TT^*)^{-1}Qy\|^2$, and one can now show in the same way as in the proof of Lemma 3.5 the validity of the assertion. □

Proof of Lemma 3.1o. A routine, albeit tedious, calculation yields that the derivative of the objective function is given by

$$[(1-K(2n-1)b_m^2/\alpha)\, f_n^m(\alpha,y)-n\delta^2]\; 2n/\alpha^2.$$

Our assertion can now be validated, using the same arguments as in the proof of Lemma 2.6. □

APPENDIX 3. COMPUTATIONAL ASPECTS.

For computing $x^n_{\alpha_m(\delta),\delta,m}$ given by (1.12), one chooses a basis $\{z_1,\dots,z_m\}$ of V_m, computes the m x m matrices $B_m = (\langle Tz_i, Tz_j\rangle)$ and $M_m := (\langle z_i, z_j\rangle)$ and the vector $y_m := (\langle Tz_i, y_\delta\rangle)$.

To compute the regularization parameter $\alpha_m(\delta)$, note that $Q_m y_\delta = \sum_{i=1}^m \mu_i Tz_i$ if and only if $\mu \in \mathbb{R}^m$ solves

(A.15) $$B_m\mu = y_m.$$

Hence,

(A.16) $$f^n_m(\alpha,y_\delta) = \alpha^{2n+1}\langle T_m(\alpha I+T^*_mT_m)^{-n}\sum_{i=1}^m \mu_i z_i,\ T_m(\alpha I+T^*_mT_m)^{-(n+1)}\sum_{i=1}^m \mu_i z_i\rangle$$

$$= (w^n)^T B_m w^{n+1},$$

where w^j is defined by

(A.17) $$\sum_{i=1}^m w^j_i\, z_i = \alpha^j(\alpha I+T^*_mT_m)^{-j}\sum_{i=1}^m \mu_i z_i.$$

For practical computations we use the iteration formula

(A.18) $$w^0 := \mu;\ (\alpha M_m+B_m)w^{j+1} = \alpha M_m w^j;\ j = 0,1,\dots,n.$$

Further, we have

$$\frac{d}{d\alpha} f^n_m(\alpha,y_\delta) = (2n+1)\alpha^{2n}\langle T_mT^*_m(\alpha I+T_mT^*_m)^{-(2n+2)}Q_m y_\delta, Q_m y_\delta\rangle$$

$$= (2n+1)\alpha^{2n}\| T^*_mT_m(\alpha I+T^*_mT_m)^{-(n+1)}\sum_{i=1}^m \mu_i z_i\|^2$$

$$= (2n+1)\alpha^{2n}\|[(\alpha I+T^*_mT_m)^{-n}-\alpha(\alpha I+T^*_mT_m)^{-(n+1)}]\sum_{i=1}^m \mu_i z_i\|^2$$

$$= (2n+1)\|\sum_{i=1}^m (w^n_i-w^{n+1}_i)z_i\|^2 = (2n+1)(w^n-w^{n+1})^T M_m(w^n-w^{n+1}).$$

Thus, we might apply a Modified Newton Method (cf. [22], Algorithm 5.4.2.4) for solving (3.4). The convergence is global and locally quadratic. Note that for $\delta = 0$ the regularization parameter is simply given by $\alpha_m(\delta) = (2n-1)Kb_m^2$.

If $\mu \in \mathbb{R}^m$ is given by (A.15), i.e., $Q_m y_\delta = \sum_{i=1}^m \mu_i T_m z_i$, we obtain

$$x^n_{\alpha,\delta,m} = \sum_{j=1}^n \alpha^{j-1}(\alpha I+T^*_mT_m)^{-j}T^*_m y_\delta = \sum_{j=1}^n \alpha^{j-1}(\alpha I+T^*_mT_m)^{-j}T^*_mT_m\sum_{i=1}^m \mu_i z_i$$

$$= \sum_{j=1}^n [\alpha^{j-1}(\alpha I+T^*_mT_m)^{-(j-1)}-\alpha^j(\alpha I+T^*_mT_m)^{-j}]\sum_{i=1}^m \mu_i z_i = \sum_{i=1}^m (\mu_i-w^n_i)z_i,$$

where w^n is given by (A.17). Thus, determination of $\alpha_m(\delta)$ yields also $x^n_{\alpha_m(\delta),\delta,m}$ without any computational effort. For the computation of b_m one can choose an orthonormal basis $\{q_1,\dots,q_{\tilde m}\}$ of $R(T_m)$, where $\tilde m := \dim R(T_m)$. Then b_m^2 is the largest eigenvalue of the $\tilde m$ x $\tilde m$ matrix

(A.19) $$C_m := (\langle (I-P_m)T^*q_i,\ (I-P_m)T^*q_j\rangle).$$

To compute the projection $(I-P_m)x$, note that $P_m x = \sum_{i=1}^m \lambda_i z_i$ if and only if

(A.2o) $\quad M_m\lambda = (\langle z_i, x\rangle)$.

In our example, we chose V_m as a space of linear splines on a uniform grid of (m+1) points in [0,1]. As basis functions we took $z_1,\ldots,z_{m+1}$ having the property that $z_i((i-1)/m) = 1$ and z_i vanishes at all other nodes. The elements of the tridiagonal matrix M_m were computed explicitly. The functions Tz_i were evaluated on a uniform grid of 4m+1 points in [0,1]. Finally, the scalar products needed for computing the elements of B_m and y_m were approximated by Milne's rule (cf. [22]). Hence, $B_m = V^TDV$, $y_m = V^TD\,\bar{y}$, where V is the (4m+1) x m matrix with elements $V_{ji} = Tz_i(j/(4m+1))$, D is a (4m+1) x (4m+1) diagonal matrix with diagonal elements representing the weights of Milne's rule, and $\bar{y} \in R^{4m+1}$ is given by $\bar{y}_j = y_\delta(j/(4m+1))$.

Note that (A.15) is equivalent to the least squares problem $\min\|D^{1/2}(V\lambda-\bar{y})\|$, which may be solved in a numerically stable way by a QR factorization with pivoting, or by a singular value decomposition of the matrix $D^{1/2}V$ (see [7, pp.162-177]). In both cases, we obtain m_1 orthogonal (with respect to the inner product in $\mathbb{R}^{4m+1}$) vectors $\tilde{q}_1,\ldots,\tilde{q}_{m_1}$ with $m_1 = \mathrm{rank}(V)$, such that $\{D^{-1/2}\tilde{q}_1,\ldots,D^{-1/2}\tilde{q}_{m_1}\}$ forms a basis for R(V) and $(D^{-1/2}\tilde{q}_i)^TD(D^{-1/2}\tilde{q}_j) = 0$ for $i \neq j$, resp. $(D^{-1/2}\tilde{q}_i)^TD(D^{-1/2}\tilde{q}_i) = 1$.

In our examples we used this set $\{D^{-1/2}\tilde{q}_1,\ldots,D^{-1/2}\tilde{q}_{m_1}\}$ as an approximation for an orthogonal basis $\{q_1,\ldots,q_{\tilde{m}}\}$ of $R(T_m)$. Using Simpson's rule, the functions T^*q_i were then approximately evaluated on a grid of (2m+1) points on [0,1].

The scalar products needed for computing the right-hand side of (A.2o), resp. the elements of C_m, are also approximated by Simpson's rule.

For our example, one could compute the numbers b_m exactly by using other integration formulas, since on each interval $[(i-1)/m, i/m]$ the function T^*Tv_j is a polynomial with maximal degree 5. However, the method described so far may also be used for other kernels.

We see that iterated Tikhonov regularization is not much more expensive to compute than ordinary Tikhonov regularization, since the iteration (A.18) always involves the same operator, i.e., one Cholesky decomposition suffices. Incidentally, computation of the matrix B_m costs more time than determining b_m and the regularization parameter $\alpha_m(\delta)$.