ANALYSIS OF ITERATED TIKHONOV: DEFECT CORRECTION FOR ILL-POSED PROBLEMS

ROSS INGRAM*, WILLIAM LAYTON†, AND NATHANIEL MAYS‡

Abstract. Given a compact operator G, we consider the ill-posed problem,

$$\text{given } y \text{ , solve } G\phi = y \text{ (approximately).}$$

Typically, in the presence of noise $y \notin Range(G)$. We consider the iterated Tikhonov method for this problem. The method selects a regularization parameter based on stability and corrects several times to increase accuracy. We show that it gives a higher accuracy approximation to the noise-free solution at each step, both in the regular and not regular cases, and that it alleviates the parameter selection difficulty of Tikhonov regularization. We also show that each update also computes the sensitivity of the Tikhonov approximation with respect to the user selected regularization parameter.

1. Introduction.

Let $G : X \to X$ (X a Hilbert space) be a compact linear operator. Given $y \in X$ we consider the problem:

$$\text{Solve } G\phi = y \text{ for } \phi \in X. \tag{1.1}$$

This problem occurs in many applications including parameter identification and image processing (e.g. the deconvolution problem from image processing, e.g., [BB98]). The problem of deconvolution of turbulent velocities also occurs in one approach to the closure problem in turbulence modeling (e.g., [BIL06], [Geu97], [AS01], [LL05]). It is well known, e.g., [AN01], [BK04], [H94], [N84], [S07], [V02], [V82], that for G compact and $Range(G)$ infinite dimensional, problem (1.1) is ill-posed. Motivated by the deconvolution problem, we shall decompose the given data y as an element of $Range(G)$ plus noise as

$$y = \overline{\phi} + \varepsilon,\ \overline{\phi} = G\phi_{true} \in Range(G) \text{ and } \varepsilon := noise. \tag{1.2}$$

Since the problem (1.1) subject to (1.2) is ill-posed, the problem thus becomes

$$\text{Given (1.2), solve (1.1) for } \phi_{true} \text{ as accurately as possible.} \tag{1.3}$$

We study herein the combination of Tikhonov regularization and defect correction for (1.3), known as iterated Tikhonov regularization. The usual Tikhonov regularized approximation to (1.3) is the solution of

$$(G^*G + \alpha I)\phi_0 := G^*y.$$

Wrapping a defect correction loop around this results in the iterated Tikhonov regularization method, [VV86]. With iterated Tikhonov, the regularization parameter α can be selected for stability (perhaps too large initially for acceptable accuracy). The approximation is then corrected several times to recover lost accuracy as follows.

*Department of Mathematics, University of Pittsburgh, Pittsburgh, PA 15260, USA, rni1@pitt.edu, http://www.math.pitt.edu/~rni1, partially supported by NSF grants DMS 0508260 and 0810385.

†Department of Mathematics, University of Pittsburgh, Pittsburgh, PA 15260, USA, wjl@pitt.edu, http://www.math.pitt.edu/~wjl, partially supported by NSF grants DMS 0508260 and 0810385.

‡Department of Mathematics, University of Pittsburgh, Pittsburgh, PA 15260, USA, nhm3@pitt.edu, http://www.pitt.edu/~nhm3, partially supported by NSF grants DMS 0508260 and 0810385.

ALGORITHM 1.1 (iterated Tikhonov). *Given data* $y = \overline{\phi} + \varepsilon$ *where* ε *is the noise in the problem data. Select* $\alpha > 0$ *and fix* J *(the number of steps)*

Solve for ϕ_0

$$(G^*G + \alpha I)\ \phi_0 = G^*\ y$$

For $j = 1, \cdots, J$ *solve for* ϕ_j

$$(G^*G + \alpha I)[\phi_j - \phi_{j-1}] = G^*(y - G\ \phi_{j-1}).$$

The difference between iterative and non-iterative regularization is not sharply defined. Often, in the former, the iteration stopping point functions as a regularization parameter, e.g., [BK04], [GG00], [HT01]. Even so, it is important to note that in defect correction methods like iterated Tikhonov J is fixed at a moderate value so it is not an iteration (properly speaking). The important question is thus asymptotic convergence as $\alpha \to 0$ and *not* as $J \to \infty$. This point is well illustrated in the experiments and amplified by the analysis in Section 4.

Even in this simple form (fixed α and J), iterated Tikhonov has great advantages over usual Tikhonov regularization. It reduces the sensitivity of the approximation error to the exact choice of the regularization parameter α and greatly increases the accuracy of the approximation (when the noise free solution is smooth), e.g., [VV86], [EHN96] Chapter 5. Interestingly, we shall also prove a superconvergence result (Theorem 2.1) that, even in the non-smooth case, accuracy is greatly increased in the noise free solution's large scales. The implementation of iterated Tikhonov involves wrapping a loop around a Tikhonov regularization routine; thus, it is also straightforward from a programming standpoint. These benefits can be enhanced further by selecting J self-adaptively and varying α at each step. These possibilities, involving extensions of the $L-$curve method, the monotone error rule and the discrepancy principle, are developed in the important work of Engl [E87], Hämarik, Palm and Raus [HPR08], Hämarik and Tautenhahn [HT01], Gfrerer [G87], Hanke and Groetsch [HG98], among others.

Each iterated Tikhonov update step requires the solution of one regularized problem. To begin, we prove that, provided the noise free solution is sufficiently regular, each step increases the asymptotic accuracy of the computed approximation, Theorems 1.3 and 1.8. This basic result is typical for defect correction type methods like iterated Tikhonov. We believe its proof is known, e.g., in the Russian literature in Vainikko and Veretennikov [VV86]. We give a proof in Section 2 for completeness and also because the proof given makes the superconvergence result follow immediately, essentially by a change of variable.

Our initial motivation was the deconvolution problem. In this problem there is a natural length scale associated with (1.3) and a natural way to define a hierarchy of scales. This idea can be abstracted in a simple way to the general problem (1.3). Within this abstraction we show in Theorem 2.1 that even in the non-regular case (in which global error estimates exhibit no advantage for iterated Tikhonov over Tikhonov) each step of iterated Tikhonov increases the accuracy of the approximation's large scales as much as predicted in the smooth case. Due to our motivation, we briefly consider the method's induced deconvolution operator / approximate inverse and show that it is self-adjoint and positive definite (Proposition 2.3). Section 3 connects iterated Tikhonov to sensitivity analysis. We show that each iterated Tikhonov step additionally computes the sensitivity (precisely defined) of the approximate solution to variations of α and uses it to increase the approximation's accuracy. By rewriting iterated Tikhonov-Lavrentiev in terms of a minimization problem in Section 4, we show how the initial steps approximate the noise free solution while if too many steps are taken the updates become a minimizing sequence for the undesired, noisy problem. Two simple numerical experiments are given in Section 5 that confirm the fundamental theoretical predictions.

1.1. Tikhonov and Tikhonov-Lavrentiev Regularization. The work horse for approximating solutions of ill-posed problems is Tikhonov regularization. It is given by selecting the regularization parameter $\alpha > 0$ and computing the approximation $\phi_0(\alpha)$ to the noise free solution by solving

$$(G^*G + \alpha I)\phi_0(\alpha) := G^*y. \tag{1.4}$$

When G is a self-adjoint and non-negative operator (Definition 1.1 below) a simpler variant of Tikhonov regularization, Tikhonov-Lavrentiev regularization[1], is applicable. It proceeds by selecting $\alpha > 0$ and computing $\phi_0(\alpha)$ now by solving

$$(G + \alpha I)\phi_0(\alpha) := y. \tag{1.5}$$

In the absence of noise, it is known, e.g., [BHTY07], [EHN96], [H95], [MP06], [V82], that the error in both (1.4) and (1.5) converges to zero at rate $\phi_{true} - \phi_0(\alpha) = O(\sqrt{\alpha})$ and at rate $\phi_{true} - \phi_0(\alpha) = O(\alpha)$ under an additional smoothness condition on the problem data. In the presence of noise, an optimal choice of α exists for which $\phi_0(\alpha)$ converges to ϕ_{true} as the size of the noise $||\varepsilon||_X \to 0$. It is also well known that an inaccurate choice of α often leads to significant loss of either stability (α chosen too small) or accuracy (α chosen too large). For a more detailed discussion of regularization methods, see the above references (among very many).

REMARK 1.2 (Tikhonov vs Tikhonov-Lavrentiev regularization). *Many of the known properties of Tikhonov regularization hold for iterated Tikhonov as well. For example, even in problems where both can be used the choice is application dependent. For example, there is often a benefit to using full Tikhonov regularization even in the self-adjoint and non-negative definite case. This is because the RHS of (1.4) is $G^*\varepsilon$ rather than ε. Often the (adjoint) operator G^* acts as a filter and $||G^*\varepsilon||_X$ is significantly smaller than $||\varepsilon||_X$. Also, in specific applications there are often advantages of regularizing by αL (with a judicious and application dependent choice of L) rather than αI.*

1.2. The Iterated Tikhonov Method. Iterated Tikhonov uncouples stability and accuracy in the choice of α. The regularization parameter α can be selected for stability (perhaps too large initially for acceptable accuracy). The approximation is then corrected several times to recover lost accuracy. The following error estimate for the J step iterated Tikhonov approximation follows from the analysis in Section 2 for the Tikhonov-Lavrentiev case.

THEOREM 1.3 (iterated Tikhonov error estimate). *Suppose*

$$\phi_{true} \in Range\left((G^*G)^J\right).$$

[1] Tikhonov regularization is Tikhonov-Lavrentiev regularization applied to the normal equations; we prove Theorem 1.8 in Section 2; Theorem 1.3 follows by replacing G by G^*G, ε by $G^*\varepsilon$ and y by G^*y.

In the absence of noise, i.e., $\varepsilon \equiv 0$, there exists a constant $C(J) < \infty$ such that the J step iterated Tikhonov error $e_J = \phi_{true} - \phi_J$ satisfies

$$||\phi - \phi_J||_X \leq C(J)\alpha^{J+1}.$$

In the presence of noise, if for some constant ε_0 we have

$$||G^* \varepsilon||_X \leq \varepsilon_0,$$

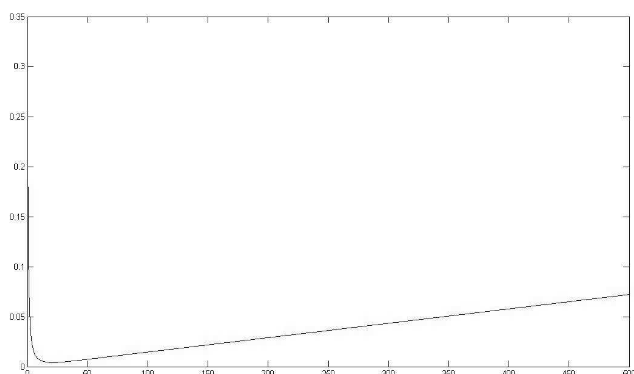
then there exists a constant $C(J) < \infty$ such that

$$\|e_J\|_X \leq \frac{(J+1)\,\varepsilon_0}{\alpha} + C\,(J)\,\alpha^{J+1}. \tag{1.6}$$

As an example (see Section 5 for details) of the increase in accuracy that the iterated Tikhonov steps give, the problem of inverting the 200×200 Hilbert matrix with relative noise of magnitude 10^{-5} is considered in the last section. For $\alpha = 0.7$ the relative error e (the scaled norm of the noise free solution - approximation) in the Tikhonov approximation (step number $j = 0$) and subsequent approximations (steps $j = 1, 2, \cdots$ and errors e) is given in the next table.

j	0	1	2	3	4	5	6	7	8	9
e	.2955	.1114	.0550	.0334	.0231	.0171	.0133	.0108	.0090	.0078
j	10	11	12	13	14	15	16	17	18	19
e	.0069	.0062	.0057	.0053	.0049	.0047	.0045	.0043	.0042	.0041
j	20	21	22	23	24	25	26	27	28	29
e	.0041	.0041	.0041	.0041	.0042	.0043	.0043	.0044	.0045	.0047

We see a clear and *dramatic* increase in accuracy in the approximation followed by an increase in the error after too many steps as predicted in Theorem 1.3, (1.6). Plotting in the next figure below the error (vertical axis) against update number (horizontal axis) in the first 500 steps shows this effect clearly.



Errors increase linearly with j after too many steps

1.3. The Iterated Tikhonov-Lavrentiev Method. Consider the case of a self-adjoint and non-negative compact operator G.

Definition 1.4. *Let $G : X \to X$. We shall say that G is* self-adjoint and non-negative definite *if $G^* = G$ and*

$$(Gv, v)_X \geq 0, \forall v \in X.$$

Tikhonov-Lavrentiev regularization applies to the self-adjoint and non-negative definite case. It is given by: select $\alpha > 0$ and solve

$$(G + \alpha I)\phi_0(\alpha) := y. \tag{1.7}$$

ALGORITHM 1.5 (iterated Tikhonov-Lavrentiev). *Let G be self-adjoint and non-negative definite. Given data $y = \overline{\phi} + \varepsilon$, select $\alpha > 0$ (the regularization parameter) and fix J (the number of steps)*

Solve for ϕ_0

$$(G + \alpha I)\ \phi_0 = \ y \tag{1.8}$$

For $j = 1, \cdots, J$ solve for ϕ_j

$$(G + \alpha I)[\phi_j - \phi_{j-1}] = y - G\ \phi_{j-1} \tag{1.9}$$

DEFINITION 1.6. *The approximate deconvolution operator / approximate inverse of G induced by the iterated Tikhonov-Lavrentiev operator is the operator $D_J : X \to X$ by*

$$D_J y = \phi_J.$$

We have the following closed form expression for the approximate inverses / approximate deconvolution operators D_J at each step.

PROPOSITION 1.7. *The deconvolution operators / approximate inverses induced by iterated Tikhonov-Lavrentiev are $D_0 = (G + \alpha I)^{-1}$, and, for $J \geq 1$,*

$$D_J = (G + \alpha I)^{-1} \sum_{j=0}^{J} \alpha^j (G + \alpha I)^{-j}.$$

Proof. These formulas follow by eliminating the intermediate steps in Algorithm 1.5. □

The basic result of iterated Tikhonov-Lavrentiev is that the approximation's accuracy increases by one power of α for each step, for example [VV86], [EHN96]. Indeed, in the absence of noise, the j^{th} step and the exact solution can both be rewritten as follows

$$\phi_j = D_0\overline{\phi} + \alpha D_0 \phi_{j-1} (j \geq 1), \ \text{ and } \quad \phi_{true} = D_0 \overline{\phi} + \alpha D_0 \phi_{true}.$$

(Recall that $D_0 = (G + \alpha I\ )^{-1}$.) Let $e_j := \phi_{true} - \phi_j$. By subtraction, the j^{th} and $(j-1)^{st}$ errors are related by

$$e_0 = \alpha D_0 \phi_{true}, \quad e_j = \alpha D_0 e_{j-1} \text{ (for } j \geq 1),$$

Thus, for $j = 0, 1, \cdots, J$,

$$e_j = \alpha^{j+1} D_0^{j+1} \phi_{true}.$$

This simple formula captures the idea of the error analysis we give. It immediately gives the preliminary result that if the solution ϕ_{true} is smooth enough that

$$\sup_{0<\alpha<1} ||D_0^{J+1} \phi_{true}||_X \leq C(J) < \infty \tag{1.10}$$

then the error is $e_J = O(\alpha^{J+1})$. Hence, each step decreases the error by one power of the regularization parameter α. The ideas in the proof of the next theorem are provided above; its detailed proof is given in Section 2.

THEOREM 1.8 (iterated Tikhonov-Lavrentiev error). *Let G be self-adjoint and non-negative definite and $\alpha > 0$. In the absence of noise, the iterated Tikhonov-Lavrentiev error $e_J = \phi_{true} - \phi_J$ satisfies*

$$e_J = \alpha^{J+1} D_0^{J+1} \phi.$$

Moreover, if $\phi_{true} \in Range\left(G^J\right)$ there exists a constant $C(J) < \infty$ such that

$$||\phi - \phi_J||_X \leq C(J)\alpha^{J+1}.$$

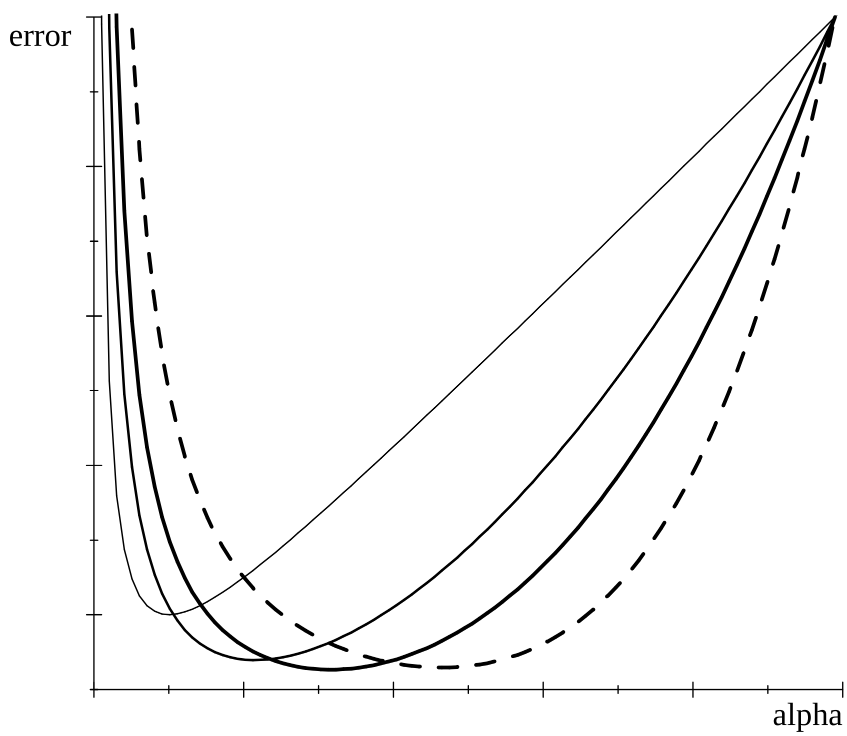
In the presence of noise if there is an ε_0 such that

$$||\varepsilon||_X \leq \varepsilon_0 < \infty \text{ and } \phi_{true} \in Range\left(G^J\right)$$

then there exists a constant $C(J) < \infty$ such that.

$$\|e_J\|_X \leq \frac{(J+1)\,\varepsilon_0}{\alpha} + C\,(J)\,\alpha^{J+1} \tag{1.11}$$

Supposing $C(J) = O(1)$, we can gain some insight can by plotting the RHS of the error estimate for various values of J in the next figure.



Errors: $J = 0$ (light),1 (medium),2 (dark),5 (dashed)

The plot suggests that as the number of steps increases

- the minimum error is obtained for a larger value of α and thus each regularized problem solved is more stable,
- the error curve is flatter near the minimum and thus the error is less sensitive to non-optimal values of α,
- the error decreases as J increases for the first steps.

1.4. Other related work. The idea of defect correction is simple and universal. In its original form, it was considered an algorithmically efficient way to perform Richardson's extrapolation [S78], [BHS84]. There is a connection between iterated Tikhonov with constant α, sensitivities and extrapolation which is explored in Section 3. Extrapolation methods based on varying α inside iterated Tikhonov is explored in depth in the work of Hämarik, Palm and Raus [HMR07] and Brezinski, Redivo-zaglia, Rodriguez and Seatzu [BRRS98]. In the context of computational fluid dynamics, the practical benefit of defect correction was recognized for nearly singular problems in the work of Hemker [Hem82a], [Hem82b] and Hemker and Koren [HK88], [HK92] with analytical proof of its effectiveness in [EL89] and [AL90]. One current view of defect correction is that it allows for approximating irregular solutions through stabilization and correction, exactly as realized in the iterated Tikhonov method. For a sample of related works outside of those cited above on ill-posed problems see, e.g., Altase and Burrage [AB94]; Axelsson and Nikolova [AN97]; Juncu [J99]; Lee [L04], Boonkkamp, Graziadei, and Mattheij [BGM04]; Heinrichs [Hei94], [Hei96]; Desideri and Hemker [DH95]; Mattheij and Nemedov [MN02]; and Crumpton and Shaw [CS93]. Another interpretation of the defect correction idea has recently been applied to integral equations in Derevtsov, Louis and Schuster [DLS04], and Schuster [S05].

2. Error analysis of iterated Tikhonov-Lavrentiev. We begin by giving a complete proof of the basic error estimate in Theorem 1.8 so that the superconvergence result (in Section 2.1) follows almost immediately. First, we construct and solve the error equation. Recall that $D_0 = (G + \alpha I)^{-1}$.

LEMMA 2.1 (error equation). *Let G be self-adjoint and non-negative definite, $\alpha > 0$. For $j = 0, 1, \cdots, J$, the j^{th} error, $e_j := \phi_{true} - \phi_j$, of the iterated Tikhonov-Lavrentiev satisfies*

$$e_0 = -D_0\varepsilon + \alpha D_0 \phi_{true}, \quad e_j = -D_0\varepsilon + \alpha D_0 e_{j-1} \text{ (for } j \geq 1\text{)}. \tag{2.1}$$

Further,

$$e_j = -D_0 \left(\sum_{i=0}^{j} \alpha^i D_0^i \right) \varepsilon + \alpha^{j+1} D_0^{j+1} \phi_{true}. \tag{2.2}$$

Proof. The iterated Tikhonov-Lavrentiev can be rewritten as

$$\phi_0 = D_0 y, \quad \phi_j = D_0 y + \alpha D_0 \phi_{j-1} \text{ (for } j \geq 1\text{)}. \tag{2.3}$$

Since $G\boldsymbol{\phi}_{true} = \overline{\phi}$ can be rearranged to read

$$\phi_{true} = D_0 \overline{\phi} + \alpha D_0 \phi_{true}, \tag{2.4}$$

subtraction yields (2.1). Eliminating intermediate quantities yields (2.2). ∎

We next show that the required regularity condition on ϕ_{true} in (1.10) holds uniformly in α if $\phi_{true} \in Range\left(G^J\right)$.

PROPOSITION 2.2 (regularity). *Let G be self-adjoint and non-negative definite. If, for integer J, $\phi_{true} \in Range\left(G^J\right)$, then there exists a constant $C(J) < \infty$ such that*

$$\sup_{0<\alpha\leq 1} \left\| D_0^J \phi_{true} \right\|_X \leq C\left(J\right). \tag{2.5}$$

More generally, if $\beta \geq 0$ and $\phi_{true} \in Range\left(G^\beta\right)$, then there exists a constant $C(\beta) < \infty$ such that

$$\sup_{0<\alpha\leq 1} \left\| D_0^\beta \phi_{true} \right\|_X \leq C\left(\beta\right).$$

Proof. For integer J, we proceed by induction. Since G is self-adjoint and non-negative definite operator, it follows that

$$inf\left\{\lambda : \lambda \in \sigma(G)\right\} \geq 0.$$

(In fact, if $Range(G)$ is infinite dimensional, it can be shown for compact G that the infimum is exactly zero.) Hence,

$$\|D_0\|_{L(X\to X)} = \left[\alpha + \inf\{\lambda \in \sigma(G)\}\right]^{-1} \leq \frac{1}{\alpha}. \tag{2.6}$$

If $\phi_{true} \in Range(G)$, then there exists an $x_1 \in X$ such that $Gx_1 = \phi_{true}$. This implies

$$\begin{aligned}
||D_0\phi_{true}||_X &= ||D_0 G x_1||_X \\
&= ||D_0[(G+\alpha I)x_1 - \alpha x_1]||_X \\
&= ||x_1 - \alpha D_0 x_1||_X \\
&\leq ||x_1||_X + \alpha ||D_0||_{L(X\to X)} ||x_1||_X \\
&\leq 2||x_1||_X =: C_1 (< \infty).
\end{aligned}$$

Now, fix $1 < j \leq J$ and suppose for any $k = 1, \cdots, j-1$ that $\phi_{true} \in Range\left(G^j\right)$ implies $\left\|D_0^k \phi_{true}\right\|_X \leq C_k$, uniformly in α. Take $\phi_{true} \in Range\left(G^j\right)$ and $x_j \in X$ with $\phi_{true} = G^j x_j$. Then, by the binomial theorem,

$$\begin{aligned}
||D_0^j \phi_{true}||_X &= ||D_0^j G^j x_j||_X \\
&= ||D_0^j[(G+\alpha I)^j x_j - P_j(\alpha, G) x_j]||_X \\
&\leq ||x_j||_X + ||D_0^j P_j(\alpha, G) x_j||_X, \text{ where} \\
P_j(\alpha, G) &= \sum_{i=1}^{j} \frac{j!}{(j-i)!i!} \alpha^i G^{j-i}.
\end{aligned}$$

Observe that $G^{j-i} x_j \in Range(G^{j-i})$ and $j - i < j$. Since $\|D_0\|_{L(X\to X)} \leq \frac{1}{\alpha}$,

$\alpha^i ||D_0^i||_{L(X\to X)} \leq 1$ so the induction hypothesis gives

$$
\begin{aligned}
||D_0^j P_j(\alpha, G)x_j||_X &\leq \sum_{i=1}^{j} \frac{j!}{(j-i)!i!}\alpha^i ||D_0^j G^{j-i} x_j||_X \\
&\leq \sum_{i=1}^{j} \frac{j!}{(j-i)!i!}\alpha^i ||D_0^i||_{L(X\to X)} ||D_0^{j-i} G^{j-i} x_j||_X \\
&\leq \sum_{i=1}^{j} \frac{j!}{(j-i)!i!} C_{j-i}.
\end{aligned}
$$

Thus,

$$
\begin{aligned}
||D_0^j \phi_{true}||_X &\leq ||x_j||_X + ||D_0^j P_j(\alpha, G)x_j||_X \\
&\leq ||x_j||_X + \sum_{i=1}^{j} \frac{j!}{(j-i)!i!} C_{j-i} =: C_j.
\end{aligned}
$$

In the non-integer case we give a different proof based on the spectral decomposition of G. Since G is a compact, non-negative, self-adjoint operator, the spectral theorem implies existence of non-negative real eigenvalues and a complete orthonormal set of eigenvectors of G:

$$
Gv_k = \lambda_k v_k, k = 1, \cdots.
$$

Denote $\widehat{\phi}(k) := (\phi_{true}, v_k)_X$ so that

$$
\phi_{true} = \sum_k \widehat{\phi}(k) v_k \text{ and } ||\phi_{true}||_X^2 = \sum_k |\widehat{\phi}(k)|^2 < \infty.
$$

Further, if $\phi_{true} \in Range\left(G^\beta\right)$ then $\phi_{true} = G^\beta \psi$ for some $\psi \in X$. Thus,

$$
||\psi||_X^2 = \sum_k |\lambda_k|^{-\beta} |\widehat{\phi}(k)|^2 < \infty.
$$

We then have by direct calculation

$$
\begin{aligned}
\left\|(G+\alpha I)^{-\beta}\phi_{true}\right\|_X^2 &= \sum_k |\lambda_k + \alpha|^{-\beta} |\widehat{\phi}(k)|^2 = \\
&= \sum_k \frac{|\lambda_k + \alpha|^{-\beta}}{|\lambda_k|^{-\beta}} \{|\lambda_k|^{-\beta}|\widehat{\phi}(k)|^2\} \leq \\
&\leq \left(\sup_{k,\alpha\geq 0} \frac{|\lambda_k + \alpha|^{-\beta}}{|\lambda_k|^{-\beta}}\right) \sum_k |\lambda_k|^{-\beta}|\widehat{\phi}(k)|^2 \leq \\
&\leq \left(\sup_{k,\alpha\geq 0} \frac{|\lambda_k + \alpha|^{-\beta}}{|\lambda_k|^{-\beta}}\right) ||\psi||_X^2.
\end{aligned}
$$

The result is thus proven provided (after rearrangement)

$$
\sup_{\lambda\in[0,||G||],\alpha\in[0,1]} \frac{|\lambda|^\beta}{|\lambda+\alpha|^\beta} \leq C < \infty.
$$

Over the indicated values the denominator is always larger than the numerator so the fraction is bounded by $C = 1$ and the result follows. $\square$

The proof of Theorem 1.8 now follows easily.

Proof. (Theorem 1.8) In the absence of noise, the claim follows by setting $\varepsilon = 0$ and iterating backwards in the error equation

$$e_j = \alpha D_0 e_{j-1} \text{ and } e_0 = \alpha D_0 \overline{\phi}.$$

This gives $e_J = \alpha^{J+1} D_0^{J+1} \overline{\phi}$. We thus have

$$||e_J||_X \leq \alpha^{J+1} ||D_0^{J+1} \overline{\phi}||_X.$$

The second claim follows by Proposition 2.2.

In the presence of noise, by taking norms across (2.2) and applying the triangle and Cauchy-Schwarz inequalities and inequality (2.6), we obtain

$$\|e_j\|_X \leq \frac{(j+1)\,\varepsilon_0}{\alpha} + \alpha^{j+1} \left\| D_0^{j+1} \phi_{true} \right\|_X. \tag{2.7}$$

Proposition 2.2 completes the proof. $\square$

To conclude this section, we prove that the iterated Tikhonov-Lavrentiev operators D_j are self-adjoint, strictly positive definite and convergent as $\alpha \to 0$.

Proposition 2.3. *Let G be bounded, self-adjoint and non-negative definite and $\alpha > 0$.*

(i) D_j is a self-adjoint, strictly positive definite and bounded operator.

(ii) Considering the noise-free problem, if Range(G^{j+1}) is dense in X, then for any $\overline{\phi} \in X$, $D_j \overline{\phi} \to \phi_{true}$ as $\alpha \to 0$ where $G\phi_{true} = \overline{\phi}$.

Proof. (i) Recall that

$$D_j = (G + \alpha I)^{-1} \sum_{k=0}^{j} \alpha^k (G + \alpha I)^{-k}.$$

D_j is a function of G hence self-adjoint. The form of this function readily shows that D_j is positive definite for $\alpha > 0$. By (2.6), $||D_j||_{L(X \to X)} \leq \frac{j+1}{\alpha}$ and hence D_j is bounded. Part (ii) follows by a density argument. $\square$

2.1. The case of less regular solutions: a superconvergence result. The assumption that $\phi_{true} \in Range\left(G^J\right)$ is a regularity condition that may or may not be satisfied. Thus two questions naturally arise:

- What is the accuracy of the iterated Tikhonov-Lavrentiev approximation in the less regular case?
- Does the approximation in the less regular case contain hidden accuracy?

We begin by noting that in the global convergence analysis (and by the same proof) less regularity yields a corresponding reduced rate of convergence.

Theorem 2.4 (Iterated Tikhonov-Lavrentiev: less regularity). *Let G be self-adjoint and non-negative definite and let $\alpha > 0$. Suppose that for some integer $\beta \geq 0$*

$$\phi_{true} \in Range\left(G^{\beta}\right) \quad \text{for some } \beta,\ 0 \leq \beta < J$$

In the absence of noise, there is a constant $C = C(J)$ such that the iterated Tikhonov-Lavrentiev error $e_J = \phi_{true} - \phi_J$ satisfies

$$||\phi_{true} - \phi_J||_X \leq C(J)\alpha^{\beta+1},\ 0 \leq \beta < J.$$

In the presence of noise, if

$$||\varepsilon||_X \leq \varepsilon_0 < \infty,$$

then there exists a constant $C(J) < \infty$ *such that*

$$\|e_J\|_X \leq \frac{(J+1)\,\varepsilon_0}{\alpha} + C\,(J)\,\alpha^{\beta+1}.$$

Some components of the error can converge faster than the global error and faster than the bound above suggests. To attach some intuition to our next result, let us begin by supposing that G can be interpreted as a smoothing operator. In that case, we can think about $\phi \in X$ as containing all scales of the solution and $G\phi \in Range(G) \subset X$ as underweighting the small scales associated with higher eigenfunctions of G. In this sense we consider $G(\phi_{true} - \phi_J)$ as representing the larger scales of the error, $G^2(\phi_{true} - \phi_J)$ as the still larger scales and so on. For example, if G represents smoothing by a Gaussian filter, there is a natural convolution length scale δ. Naturally, $G(\phi_{true} - \phi_J)$ represents the $O(\delta)$ and larger scales of the error, $G^2(\phi_{true} - \phi_J)$ the $O(\sqrt{2}\delta)$ and larger scales and, in general, $G^J(\phi_{true} - \phi_J)$ the $O(\sqrt{J}\delta)$ and larger scales.

THEOREM 2.5 (superconvergence of iterated Tikhonov). *Let G be self-adjoint and non-negative definite, let $\alpha > 0$ and suppose $\phi_{true} \in X$. If $\left\|G^J\varepsilon\right\|_X \leq \varepsilon_0$, then there exists a constant $C(J) < \infty$ such that*

$$\left\|G^J e_J\right\|_X \leq \frac{(J+1)\,\varepsilon_0}{\alpha} + C\,(J)\,\alpha^{J+1}.$$

More generally, if $\phi_{true} \in Range\left(G^\beta\right)$ for some β, $0 \leq \beta < J$ and $||G^{J-\beta}\varepsilon||_X \leq \varepsilon_0$, then

$$\left\|G^{J-\beta} e_J\right\|_X \leq \frac{(J+1)\,\varepsilon_0}{\alpha} + C\,(J)\,\alpha^{J+1}.$$

Proof. This proof follows quickly from the analysis in the proof of Theorem 1.8 by a change of variables. Indeed, define

$$\widetilde{\phi} = G^{J-\beta}\phi \text{ (in all cases with all subscripts)}, \widetilde{y} = G^{J-\beta}y, \widetilde{\varepsilon} = G^{J-\beta}\varepsilon.$$

Multiplying $G\phi = y$ through by $G^{J-\beta}$ gives a problem of exactly the same form as (1.3) with tildes over all variables:

$$\text{Solve } G\widetilde{\phi} = \widetilde{y} \text{ for } \widetilde{\phi}_{true} \text{ as accurately as possible.}$$

for $\widetilde{\phi}_{true}$ as accurately as possible. If $\phi_{true} \in Range(G^\beta)$ then clearly $\widetilde{\phi}_{true} \in Range(G^J)$. Similarly, multiplying the iterated Tikhonov-Lavrentiev algorithm by $G^{J-\beta}$ shows that the approximations $\widetilde{\phi}_j = G^{J-\beta}\phi_j$ are the iterated Tikhonov-Lavrentiev approximations to the tilde problem. The basic error estimate for $\widetilde{e}_J = \widetilde{\phi}_{true} - \widetilde{\phi}_J$ now applies to the tilde problem. Since $\widetilde{e}_J = G^{J-\beta}e_J$ the result follows. ∎

3. Sensitivity of Tikhonov-Lavrentiev regularization. Sensitivities give important information about the reliability of predictions, e.g., [G02], [AL07], [SS02], [BB97], [L94]. They are also required when the output of an algorithm is optimized over the algorithm's inputs, [G02].

DEFINITION 3.1. *The sensitivity with respect to* α*,* $s_j(\alpha)$*, of the* j^{th} *iterated Tikhonov-Lavrentiev approximation* ϕ_j *is*

$$s_j(\alpha) := \frac{d}{d\alpha}\phi_j(\alpha).$$

Sensitivity equations are obtained by implicitly differentiating the equations derived for the proposed algorithms. In the initial Tikhonov-Lavrentiev approximation, implicit differentiation gives the following coupled system for $\phi_0(\alpha)$ and $s_0(\alpha)$:

$$\begin{aligned}(G+\alpha I)\phi_0(\alpha) &= y \\ (G+\alpha I)s_0(\alpha) &= -\phi_0(\alpha)\end{aligned} \tag{3.1}$$

Thus, calculating the sensitivity of ϕ_0 involves one extra regularization solve step beyond the Tikhonov-Lavrentiev regularization solve. Interestingly, the extra work computing α sensitivities of the Tikhonov-Lavrentiev approximation is not wasted from the standpoint of accuracy. Indeed, following an idea in [AL07], the sensitivity can be used to increase the accuracy of the Tikhonov-Lavrentiev approximation as follows. In the noise-free problem, since $\phi_{true} = \phi_0(\alpha)|_{\alpha=0}$, one Newton step in α parameter space gives the approximation $\phi_0(\alpha) - \alpha s_0(\alpha)$ which has accuracy

$$\phi_{true} = \phi_0(\alpha) - \alpha s_0(\alpha) + O(\alpha^2). \tag{3.2}$$

We begin by showing that this sensitivity corrected approximation $\phi_0(\alpha) - \alpha s_0(\alpha)$ in (3.2) coincides with the $J = 1$ step of the iterated Tikhonov-Lavrentiev algorithm.

PROPOSITION 3.2. *(i) The sensitivity* $s_0(\alpha)$ *satisfies*

$$s_0(\alpha) = \frac{\phi_0(\alpha) - \phi_1(\alpha)}{\alpha}. \tag{3.3}$$

(ii) The $J = 1$ *iterated Tikhonov-Lavrentiev approximation is identical to the sensitivity corrected approximation (3.2):*

$$\phi_1 \equiv \phi_0(\alpha) - \alpha s_0(\alpha). \tag{3.4}$$

Proof. Starting with $(G+\alpha I)\phi_0 = y$ and $(G+\alpha I)(\phi_1 - \phi_0) = y - G\phi_0$, the result follows by implicit differentiation and algebraic rearrangement □

The correspondence between iterated Tikhonov-Lavrentiev and sensitivity corrections of Tikhonov-Lavrentiev regularization can be continued to higher sensitivities and more than one update step. Indeed, if $\phi_0(\alpha)$ is a sufficiently regular function of α, we can develop its Taylor polynomial at α as a function of $\widetilde{\alpha}$ as

$$T_k(\phi_0(\alpha))(\widetilde{\alpha}) := \phi_0(\alpha) + (\widetilde{\alpha} - \alpha)\frac{d\phi_0(\alpha)}{d\alpha} + ... + \frac{(\widetilde{\alpha} - \alpha)^k}{k!}\frac{d^k\phi_0(\alpha)}{d\alpha^k}. \tag{3.5}$$

Beginning with this, we can take $T_k(\phi_0(\alpha))(\widetilde{\alpha})|_{\widetilde{\alpha}=0}$ as an approximation to ϕ_{true}, aiming at order of accuracy $O(\alpha^{k+1})$. This gives

$$\begin{aligned}Approx_k(\phi_{true}) \ : &= T_k(\phi_0(\alpha))(\widetilde{\alpha})|_{\widetilde{\alpha}=0} \\ &= \phi_0(\alpha) - \alpha\frac{d\phi_0(\alpha)}{d\alpha} + ... + \alpha^k\frac{(-1)^k}{k!}\frac{d^k\phi_0(\alpha)}{d\alpha^k}.\end{aligned} \tag{3.6}$$

We show next that (3.6) is exactly the k^{th} approximation and thus the updates implicitly compute higher sensitivities of $\phi_0(\alpha)$ and use them to correct the approximation.

THEOREM 3.3. *Let G be self-adjoint and non-negative definite and $\alpha > 0$. Consider the higher order, sensitivity corrected approximation (3.6). In the absence of noise and for sufficiently regular $\phi_0(\alpha)$*

$$\phi_{true} = Approx_k(\phi_{true}) + O(\alpha^{k+1}).$$

Further, the k^{th} iterated Tikhonov-Lavrentiev approximation $\phi_k(\alpha)$ is exactly (3.6):

$$\phi_k(\alpha) = Approx_k(\phi_{true}).$$

We begin the proof with several algebraic identities.

LEMMA 3.4. *(i) For $j > 0$,*

$$(G + \alpha I)[s_j(\alpha) - s_{j-1}(\alpha)] = -(\phi_j(\alpha) - \phi_{j-1}(\alpha)) - Gs_{j-1}(\alpha).$$

(ii) For $j > 1$,

$$s_{j-1}(\alpha) = \frac{\phi_{j-1}(\alpha) - \phi_j(\alpha)}{\alpha} + \alpha D_0 s_{j-2}(\alpha).$$

(iii) For $j > 1$,

$$\phi_j(\alpha) = \phi_{j-1}(\alpha) - \alpha s_{j-1}(\alpha) + \alpha^2 D_0 s_{j-2}(\alpha).$$

(iv) For $j > 0$,

$$(G + \alpha I)^{j-1}(\phi_j(\alpha) - \phi_{j-1}(\alpha)) = \alpha^{j-1}(\phi_1(\alpha) - \phi_0(\alpha))$$

Proof. Part (i) Follows directly from implicit differentiation with respect to α of equations (1.8) and (1.9). Part (ii) follows from (i) and equations (1.8) and (1.9). Indeed,

$$(G + \alpha I)(s_{j-1}(\alpha) - s_{j-2}(\alpha)) = -(\phi_{j-1}(\alpha) - \phi_{j-2}(\alpha)) - Gs_{j-2}(\alpha)$$

implies that

$$(G + \alpha I)s_{j-1}(\alpha) = -(\phi_{j-1}(\alpha) - \phi_{j-2}(\alpha)) + \alpha s_{j-2}(\alpha).$$

Further

$$(G + \alpha I)(\phi_{j-1}(\alpha) - \phi_{j-2}(\alpha)) = y - G\phi_{j-2}(\alpha)$$

implies as well that

$$-\alpha\phi_{j-2}(\alpha) = y - (G + \alpha I)\phi_{j-1}(\alpha), \text{ and} \tag{3.7}$$
$$-\alpha\phi_{j-1}(\alpha) = y - (G + \alpha I)\phi_j(\alpha). \tag{3.8}$$

So, we conclude that,

$$\alpha(G + \alpha I)s_{j-1}(\alpha) = -(G + \alpha I)(\phi_j(\alpha) - \phi_{j-1}(\alpha)) + \alpha^2 s_{j-2}(\alpha)$$

The desired conclusion follows immediately. Part (iii) follows immediately by algebraically rearranging (ii). For part (iv) note that in the proof of (ii) we showed that

$$(G+\alpha I)\phi_{j-1}(\alpha) = y + \alpha\phi_{j-2}(\alpha) \tag{3.9}$$

$$(G+\alpha I)\phi_j(\alpha) = y + \alpha\phi_{j-1}(\alpha). \tag{3.10}$$

So, we have

$$(G+\alpha I)(\phi_j(\alpha) - \phi_{j-1}(\alpha)) = \alpha(\phi_{j-1}(\alpha) - \phi_{j-2}(\alpha)),$$

from which the conclusion follows. □

LEMMA 3.5. *(i) We have*

$$\frac{ds_0(\alpha)}{d\alpha} = -\frac{s_1(\alpha)}{\alpha}.$$

(ii) For $k = 2, \cdots, J$

$$\frac{d^{k-1}s_0(\alpha)}{d\alpha^{k-1}} = (-1)^{k-1}\frac{k!}{2\alpha}D_0^{k-2}s_1(\alpha).$$

(iii) The sensitivity $s_1(\alpha)$ of the first iterated Tikhonov-Lavrentiev approximation $\phi_1(\alpha)$ satisfies

$$(G+\alpha I)s_1(\alpha) = -2(\phi_1(\alpha) - \phi_0(\alpha))$$

Proof. Part (i) follows by implicit differentiation of equation (3.3) with respect to α. For part (ii), starting with $(G+\alpha I)\phi_0(\alpha) = y$ and differentiating k times with respect to α we get

$$\frac{d^{k-1}s_0(\alpha)}{d\alpha^{k-1}} = -kD_0\frac{d^{k-2}s_0(\alpha)}{d\alpha^{k-2}}$$

By continued substitution and using $\frac{ds_0(\alpha)}{d\alpha} = -\frac{s_1(\alpha)}{\alpha}$ in from (i) in the final step, we get (ii). For (iii), $s_1(\alpha)$ satisfies

$$(G+\alpha I)[s_1(\alpha) - s_0(\alpha)] = -(\phi_1(\alpha) - \phi_0(\alpha)) - Gs_0(\alpha).$$

Rearranging and inserting $s_0(\alpha) = -\frac{\phi_1(\alpha)-\phi_0(\alpha)}{\alpha}$ gives (iii). □

We are now ready to prove Theorem 3.3.

Proof. (Theorem 3.3) We first note that for $\alpha > 0$, $\phi_0(\alpha) = (G+\alpha I)^{-1}y$ is a smooth function of α. By Taylor's theorem, for regular $\phi_0(\alpha)$,

$$\phi_{true} = T_k(\phi_0(\alpha)) + O(\alpha^{k+1}).$$

Next, by Lemma 3.4, part (iii) we have

$$\phi_k(\alpha) = \phi_{k-1}(\alpha) - \alpha s_{k-1}(\alpha) + \alpha^2 D_0 s_{k-2}(\alpha).$$

We want to show that this is exactly the k^{th} term in the claimed Taylor expansion. We have already shown this is true for $k = 1$. We assume that the claim is true for $k-1$ and proceed by induction. Hence, we must show

$$\alpha^k\frac{(-1)^k}{k!}\frac{d^k\phi_0(\alpha)}{d\alpha^k} = \alpha s_{k-1}(\alpha) + \alpha^2 D_0 s_{k-2}(\alpha)$$

which, applying Lemmas (3.4) (iii) and (3.5) (ii), is equivalent to showing

$$\phi_k(\alpha) - \phi_{k-1}(\alpha) = -\alpha^{k-1}\frac{1}{2}D_0^{k-2}s_1(\alpha),$$

which, multiplying by $-\frac{2}{\alpha^{k-1}}(G+\alpha)^{k-1}$, is equivalent to

$$(G+\alpha I)s_1(\alpha) = -\frac{2}{\alpha^{k-1}}(G+\alpha I)^{k-1}(\phi_k(\alpha) - \phi_{k-1}(\alpha))$$

which, by Lemma 3.4, part (iv), is equivalent to

$$(G+\alpha I)s_1(\alpha) = -2(\phi_1(\alpha) - \phi_0(\alpha)).$$

Lemma 3.5, part (iii) asserts the validity of the last equation. By equivalence, we get the desired conclusion. □

We offer an equivalent algorithm for iterated Tikhonov-Lavrentiev using the sensitivities s_j developed above. If sensitivities are desired, this reformulation algorithm is an efficient way to compute s_j and update ϕ_{j+1} as needed.

ALGORITHM 3.6 (iterated Tikhonov-Lavrentiev in terms of sensitivities). *Let G be self-adjoint and non-negative definite. Given data $y = \overline{\phi} + \varepsilon$, select $\alpha > 0$ and fix J*

Solve for ϕ_0 and s_0,

$$(G+\alpha I)\ \phi_0 = \ y$$
$$(G+\alpha I)\ s_0 = -\phi_0.$$

Set

$$\phi_1 = \phi_0 - \alpha s_0.$$

For $j = 1, \cdots, J$

Solve for s_j

$$(G+\alpha I)[s_j - s_{j-1}] = -(\phi_j - \phi_{j-1}) - Gs_{j-1}$$

If $j \neq J$ then

$$\phi_{j+1} = \phi_j - \alpha s_j + \alpha^2 D_0 s_{j-1}$$

4. Descent properties of iterated Tikhonov-Lavrentiev approximations. The original problem (1.3) for self-adjoint and non-negative definite G is formally equivalent to the minimization problem

$$\text{minimize}_{v\in X}\ J_\varepsilon(v)\ \text{, where } J_\varepsilon(v) := \frac{1}{2}(Gv, v)_X - (\overline{\phi} + \varepsilon, v)_X.$$

The question naturally arises if the approximations $\phi_0, \phi_1, \cdots$ form a minimizing sequence for $J_\varepsilon(\cdot)$ and/or the noise - free functional $J_0(\cdot) := J_\varepsilon(\cdot)|_{\varepsilon\equiv 0}$.

PROPOSITION 4.1. *Let G be self-adjoint and non-negative definite and $\alpha > 0$. Then the the iterated Tikhonov-Lavrentiev iterates are a minimizing sequence for J_ε. In particular,*

$$J_\varepsilon(\phi_j) - J_\varepsilon(\phi_{j+1}) = \frac{1}{2}((G+2\alpha I)(\phi_{j+1} - \phi_j), \phi_{j+1} - \phi_j) \geq 0. \qquad (4.1)$$

Thus

$$J_\varepsilon(\phi_{j+1}) < J_\varepsilon(\phi_j), \text{ unless } \phi_{j+1} = \phi_j.$$

Proof. The formula is an identity, well-known in numerical linear algebra from the work of Householder [Hou06]. It follows here as well by expanding both sides and cancelling terms. Formula (4.1) immediately implies $J_\varepsilon(\phi_{j+1}) < J_\varepsilon(\phi_j)$, unless $\phi_{j+1} = \phi_j$ as claimed. □

From the form of the updates, $\phi_{j+1} = \phi_j$ only if $G\phi_j = \overline{\phi} + \varepsilon$, see (1.9). Thus, if one continues to update, as $j \to \infty$, ϕ_j converges to the undesired solution of the noisy data problem, reinforcing the idea that it is critical to stop after a few update steps and not iterate.

Since the problem we seek to solve is instead the noise-free one

$$G\phi_{true} = \overline{\phi},$$

it is also natural to ask under what conditions the updates reduce the noise-free functional

$$J_0(v) := \frac{1}{2}(Gv, v)_X - (\overline{\phi}, v)_X.$$

Rearranging the formula (4.1) shows that

$$J_0(\phi_j) - J_0(\phi_{j+1}) = (\varepsilon, \phi_{j+1} - \phi_j) + \frac{1}{2}((G + 2\alpha I)(\phi_{j+1} - \phi_j), \phi_{j+1} - \phi_j). \quad (4.2)$$

THEOREM 4.2. *Let G be self-adjoint and non-negative definite and $\alpha > 0$. Suppose the estimate on the noise $\varepsilon_0 \geq ||\varepsilon||_X$ is known. Then the the iterated Tikhonov-Lavrentiev approximations are a minimizing sequence for the noise-free functional J_0 as long as*

$$\alpha \geq \frac{\varepsilon_0}{||\phi_{j+1} - \phi_j||_X}. \quad (4.3)$$

Proof. This follows from (4.2) and the Cauchy-Schwarz inequality. □

Theorem 4.2 suggests that in the early steps, in which larger updates are expected, the updates move the approximate solution closer to the noise free solution. Later in the process, as the updates become smaller, ϕ_j begins to deviate from an approximation of the noise free solution unless α is increased. If more is known about the noise or its statistical distribution, (e.g., if there is a projection operator $\overline{P}$ with $\overline{P}\varepsilon = 0$) then

$$(\varepsilon, \phi_{j+1} - \phi_j)_X = (\varepsilon, (I - \overline{P})[\phi_{j+1} - \phi_j])_X.$$

In other words, if the component of the approximation in $Range(\overline{P})$ is updated the approximation is still one to the noise free problem. This suggests the following small algorithmic modification.

ALGORITHM 4.3 (modified Tikhonov-Lavrentiev DCM). *Given data $\overline{\phi} + \varepsilon$, Suppose $\varepsilon_0 \geq ||\varepsilon||_X$ and given $\overline{P}$ with $\overline{P}\varepsilon = 0$. Select α and fix J .*

Solve for ϕ_0

$$(G + \alpha I)\ \phi_0 = \overline{\phi}.$$

For $j = 1, \cdots, J$ *and while* $\alpha \geq \frac{\varepsilon_0}{||\phi_j - \phi_{j-1}||_X}$ *solve for* ϕ_j

$$(G + \alpha I)[\phi_j - \phi_{j-1}] = \overline{\phi} - G\ \phi_{j-1}$$

If $\alpha < \frac{\varepsilon_0}{||\phi_j - \phi_{j-1}||_X}$ *then either increase* α *so that (4.3) holds and recompute or compute as above* $\phi_j - \phi_{j-1}$ *and set*

$$\phi_j \Leftarrow \phi_{j-1} + \overline{P}(\phi_j - \ \phi_{j-1})$$

Set $D\overline{\phi} := \phi_J$.

5. Numerical Illustrations. We consider two problems illustrating the iterated Tikhonov and iterated Tikhonov-Lavrentiev methods: inverting the Hilbert matrix with a noisy right hand side and the inverse Laplace transform with noisy data.

Test 1: iterated Tikhonov-Lavrentiev for the 200 × 200 Hilbert matrix. The $N \times N$ Hilbert matrix, H_N , has ij entries given by

$$(H_N)_{ij} := \frac{1}{i + j - 1}, 1 \leq i, j \leq N.$$

The Hilbert matrix is well known to be extremely ill-conditioned, e.g., [S06], and has been used for testing methods for solving ill–posed problems, in, for example, [LR08]. It is known that H_{200} has condition number

$$cond(H_{200}) \simeq O(\frac{e^{3.5255 \times 200}}{\sqrt{200}}) = 1.1763e + 305.$$

We take the exact solution to the noise free problem to be

$$\phi_{true} = [1, \frac{1}{2}, \cdots, \frac{1}{200}]$$

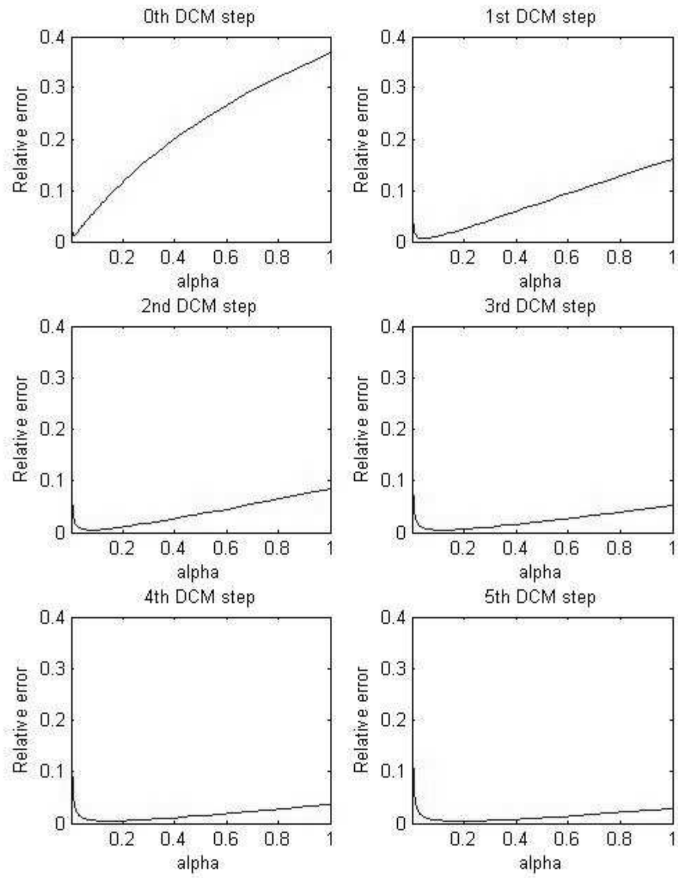
and compute the noise free RHS in extended precision. To this we add noise given by

$$\varepsilon := unit(rand(200)) \cdot 10^{-4} \cdot \|y\|, \text{ so that } \frac{||\varepsilon||}{||y||} = 10^{-4},$$

where $unit(rand(200))$ is the Matlab routine used to generate a random unit vector and $|| \cdot ||$ is the usual euclidean norm. H_N is self-adjoint and positive definite, so we applied iterated Tikhonov-Lavrentiev with parameters α and J ranging over values

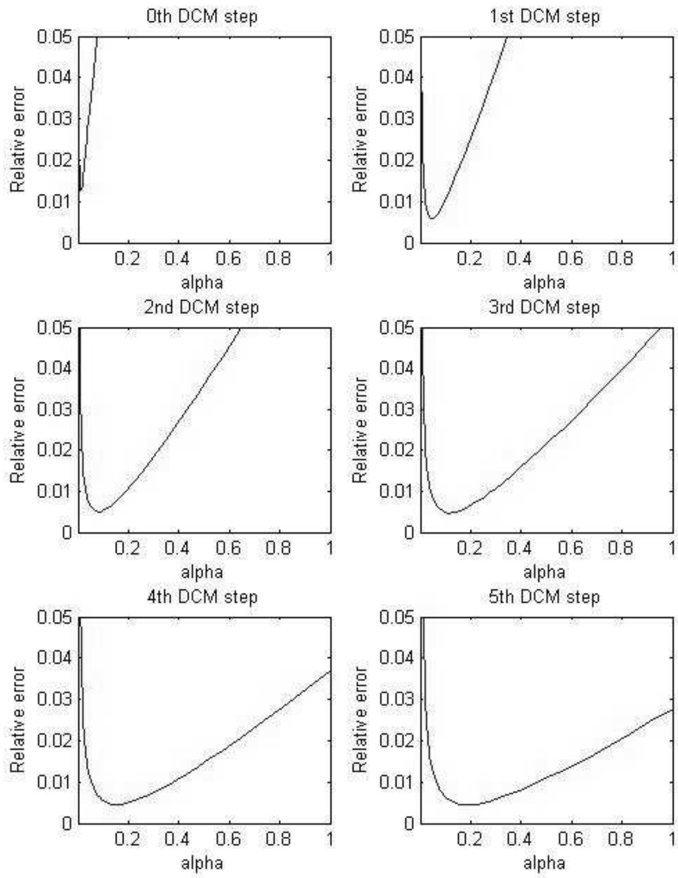
$$10^{-4} \leq \alpha \leq 1, \quad \text{and} \quad J = 0, 1, \cdots, 5.$$

For the range of α tested, we consistently observed a dramatic increase in accuracy with the updated approximations; see the error table in the introduction for a representative case. The next figure plots the error between the true, noise-free solution and first 5 iterated Tikhonov-Lavrentiev approximations for the 200 × 200 Hilbert matrix.



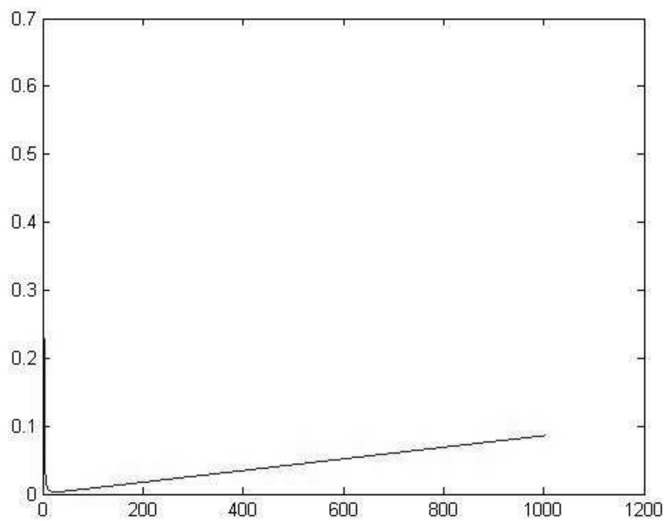
Relative error vs. α for 5 steps for H_{200}

The horizontal axis is the range of regularization parameter values: α from 10^{-4} to 1. The vertical axis plots the relative error between the calculated solution with the indicated update's approximation and the noise free solution. The error plots above in the last figure show an increase in accuracy with the updates and that the error becomes less sensitive to the regularization parameter as J increases. (The region of the error curve around the minimum becomes flatter as J increases.) This is also in accordance with the theoretical predictions. (Compare this figure with the corresponding figure illustrating the theorem's predictions in the introduction.) The next figure zooms in to the part of the relative error occurring in the range from 0 to 0.05 for a clearer view of these effects.



Zoom of: Relative error vs. α, 5 DCM steps for H_{200}

With any update method, the natural question arises of what happens when the iteration proceeds and $J \to \infty$. (We stress that in practical computing one should stop at a moderate value of J, e.g., $J = 5$.) The theory predicts that as $J \to \infty$ the updates' error first decreases then increases as the updates converge to the exact solution of the noisy problem. This is observed in the calculations as well. The next figure plots the update steps from $J = 0$ to $J = 1000$ (a ridiculously large number of update steps) for a fixed $\alpha = 0.5$.



Absolute error vs. J , $\alpha = 0.5$ for H_{100}: Don't Iterate!

This first test shows that we can calculate the noise-free solutions of problems with noise with good accuracy and with little prior knowledge of the regularization parameter.

Test 2: iterated Tikhonov for inverting the Laplace transform.

The second example we look at is an inverse Laplace Transform problem from [LR08]. The discretized inverse-Laplace transform matrix is well known to be extremely ill-conditioned, e.g., [H94], and has been used as a problem in [LR08] for testing methods for solving ill–posed problems. The data is selected so that the exact, noise free solution is given by

$$\int_0^\infty e^{-s\cdot t} x(s) ds = b(t), t > 0 \text{ where}$$
$$b(t) = \frac{1}{t + \frac{1}{2}} \quad \text{and} \quad x(s) = e^{-t/2}.$$

The problem is converted into a discrete system using a 100 point Gauss-Laguerre quadrature with 100 equidistantly spaced points $0 \le s(i) \le 10$. The discretization parameters $t(j)$ come from solving the eigensystem derived from symmetric tridiagonal recurrence relation for Laguerre polynomials. Also, we choose for $1 \le i \le 100$,

$$s(i) = i/10, \;\; x(i) = x(s(i)) \text{ and } b(i) = b(t(i)).$$

This discretization gives a matrix $A \in \mathbb{R}^{100\times 100}$ that is neither symmetric nor non-negative. Thus, this illustration tests the full iterated Tikhonov algorithm with parameters ranging over values

$$10^{-6} \le \alpha \le 1 \text{ and } \;\; J = 0, 1, \cdots, 5,$$

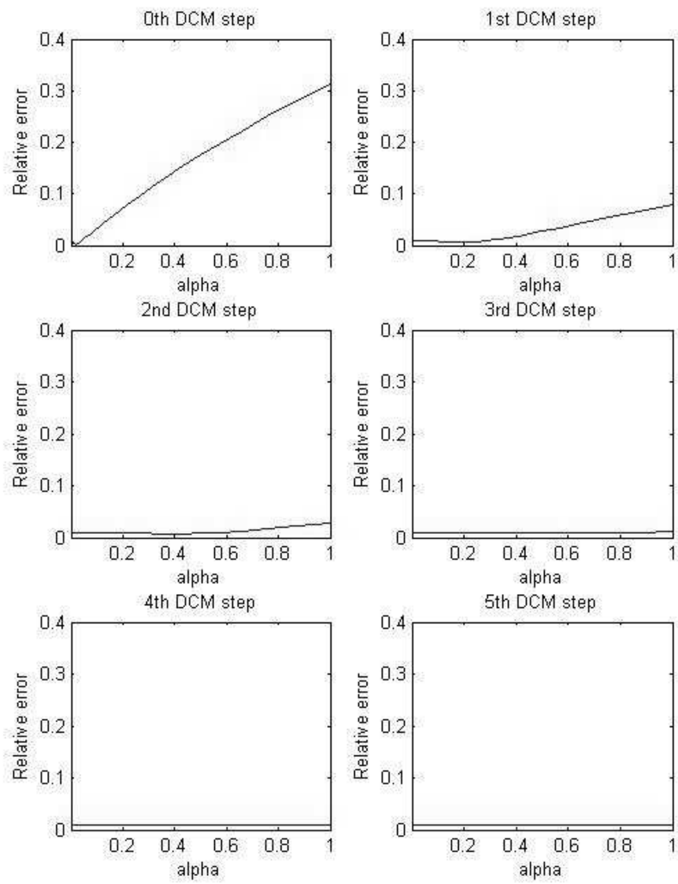
with noise introduced of magnitude

$$\|\varepsilon\| = 10^{-2} \cdot \|b(j)\| \;\; \text{with} \;\; \frac{||\varepsilon||}{||y||} = 10^{-2}.$$

The 100×100 Inverse-Laplace matrix is known to be highly ill-conditioned (see [H94]). We take the exact solution to the noise free problem to be

$$\phi_{true} = e^{t/2}$$

and compute the noise free RHS in extended precision. The next figure plots the relative error in the Tikhonov approximation ($J = 0$) and the first 5 steps (i.e., $J = 1, \cdots, 5$)applied to the above 100×100 inverse Laplace example.



Relative error vs. α, 5 steps Tikhonov DCM: inverse Laplace transform

The horizontal axis is the regularization parameter α, from $10^{-6} \leq \alpha \leq 1$. The relative noise has magnitude 10^{-2}. The vertical axis shows the relative error between the approximation and the noise-free solution. In this graph the error becomes less sensitive to the chosen values of the regularization parameter α as J increases. This observation and the error behavior is in accordance with the theoretical predictions. The second test show that iterated Tikhonov is a very successful way to calculate solutions to ill-posed, non-self-adjoint and indefinite problems accurately.

6. Conclusions. Defect correction in the form of the iterated Tikhonov and iterated Tikhonov-Lavrentiev methods is a powerful tool for increasing accuracy of regularization methods for ill-posed problems. While optimal parameters also exist for these methods, one important feature of them is that they do not require optimal parameters for accuracy. Iterated Tikhonov reduces the sensitivity of the error to the exact value of the regularization parameter selected. They are also very efficient (in programmer effort) to implement once a basic regularization method is available. It is almost a meta-theorem in programming that simpler is better and in numerical analysis that higher accuracy methods give better approximations (even when not predicted by theory). Since these two conditions are often exclusive, it is particularly interesting to note that iterated Tikhonov and iterated Tikhonov-Lavrentiev are both algorithmically simple and highly accurate, even in the less regular case!

REFERENCES

[AS01] N. A. Adams and S. Stolz, *Deconvolution methods for subgrid-scale approximation in large eddy simulation*, Modern Simulation Strategies for Turbulent Flow, R.T. Edwards, 2001.

[AN01] A.K. Alekseev and I.M. Navon, *The analysis of an ill-posed problem using multi-scale resolution and second order adjoint techniques*, C.M.A.M.E. 190(2001) 1486-1496.

[AB94] I. Altas and K. Burrage, *A high-accuracy defect- correction multigrid method for the steady incompressible Navier-Stokes equations*, J. Comput. Phys 114(2) (1994), 227-233.

[AL07] M. Anitescu and W. Layton, *Sensitivities in large eddy simulation and improved estimates of turbulent flow functionals*, SIAM J. Scientific Computing 29(2007) 1650-1667.

[AL90] O. Axelsson and W Layton, *Defect correction methods for convection-dominated, convection diffusion equations*, RAIRO J. Numer. Anal., 24(1990) 423-455.

[AN97] O. Axelsson and M. Nikolova, *Adaptive refinement for convection-diffusion problems based on a defect-correction technique and finite difference method*, Computing 58(1) (1997), 1-30.

[BK04] A.B. Bakushinsky, and M. Yu. Kokurin, *Iterative methods for approximate solution of inverse problems*, Kluwer, Dordrecht. the Netherlands, 2004.

[BIL06] L. C. Berselli, T. Iliescu, and W. Layton, *Mathematics of Large Eddy Simulation of Turbulent Flows*. Springer, Berlin, 2006.

[BB98] M. Bertero and B. Boccacci, *Introduction to Inverse Problems in Imaging*, IOP Publishing Ltd., 1998.

[BHS84] K. Böhmer, P. Hemker, H. Stetter, *The defect correction approach*, in: K. Böhmer, H. Stetter (Eds.), Defect Correction Methods: Theory and Applications, Springer, Berlin, 1984, pp. 1-32.

[BGM04] J.H.M.T. Boonkkamp, M. Graziadei, R.M.M. Mattheij, *Local defect correction with slanting grids*, Numer. Meth. Part. Diff. Eqns. 20(1) (2004), 1-17.

[BB97] J. Borggaard and J. Burns, *A PDE sensitivity equation method for optimal aerodynamic design*, J.C.P. 136(1997), 366-384.

[BHTY07] A. Böttcher, B. Hoffman, U. Tautenhahn and M. Yamamoto, *Convergence rates for Tikhonov regularization from different kinds of smoothness*, Appl. Anal. 85(2006) 555-578.

[BRRS98] C. Brezinski, M. Redivo-zaglia, G. Rodriguez and S. Seatzu, *Extrapolation techniques for ill-conditioned linear systems*, Numer. Math. 81(1998) 1-29.

[CS93] P.I. Crumpton and G.J. Shaw, *An analysis of the finite termination property of the defect correction method*, Int. J. Numer. Meth. Fluids 16(3) (1993), 199-215.

[DLS04] E. Derevtsov, A. Louis and T. Schuster, *Two approaches to the problem of defect correction in vector field tomography solving boundary value problems*, J. Inv. Ill-Posed Problems, 12 (2004) 597-626.

[DH95] J.A. Desideri and P.W. Hemker, *Convergence analysis of the defect-correction iteration for hyperbolic problems*, SIAM J. Sci. Comput. 16(1) (1995), 88-118.

[E87] H.W. Engl, *On the choice of the regularization parameter for iterated Tikhonov-regularization of inverse problems,* J. Approx. Theory 49(1987) 55-63.

[EHN96] H.W. Engl, M. Hanke and A. Neubauer, *Regularization of inverse problems,* Kluwer, Dordrecht, the Netherlands, 1996.

[EL89] V Ervin and W Layton, *A study of defect correction, finite difference methods for* *convection diffusion equations*, SIAM J.N.A. 26(1989), 169-179.

[Geu97] B. J. Geurts, *Inverse modeling for large eddy simulation*, Phys. Fluids, 9(1997), 3585.

[G03] B. J. Geurts, *Elements of direct and large eddy simulation,*Edwards Publishing, 2003.

[G87] H. Gfrerer, *An a posteriori parameter choice for ordinary and iterated Tikhonov regularization leading to optimal convergence rates*, Math. Comp. 49(1987) 507-522.

[GG00] S.F. Gilyazov and N.L. Goldman, *Regularization of ill-posed problems by iteration methods*, Kluwer, Dordrecht, 2000.

[G02] M Gunzburger, *Sensitivities, adjoints and flow optimization*, vol 15 in: Advanced Design Controls, SIAM, Philadelphia, 2002.

[HMR07] U. Hämarik, R. Palm and T. Raus, *Use of extrapolation in regularization methods*, J. Inverse and Ill-Posed Problems, 15(2007) 277-294.

[HPR08] U. Hämarik, R. Palm and T. Raus, *Extrapolation of Tikhonov and Lavrentiev regularization methods*, technical report 2008.

[HT01] U. Hämarik and U. Tautenhahn, *On the monotone error rule for parameter choice in iterative and continuous regularization methods,* BIT 41(2001) 1029-1038.

[HG98] M. Hanke and C.W. Groetsch, *Nonstationary iterated Tikhonov regularization*, J. Optim. Theory and Appl. 98(1998) 37-53.

[H95] M. Hegland, *Hilbert scales and their interpolation inequalities with applications to Tikhonov regularization*, Appl. Anal. 59(1995)207-223.

[H92] P. C. Hansen, *Analysis of discrete ill-posed problems by means of the L-curve*, SIAM Review 34(1992) 561-580.

[H94] P. C. Hansen, *Regularization tools: A Matlab package for analysis and solution of discrete ill-posed problems*, Numer. Algorithms, 6 (1994), pp. 1–35. (Software available at http://www.netlib.org.)

[Hei94] W. Heinrichs, *Defect correction for the advection-diffusion equation*, Comput. Method App. M. 119(3-4) (1994), 191-197.

[Hei96] W. Heinrichs, *Defect correction for convection dominated flow*, SIAM J. Sci. Comput. 17(5) (1996), 1082-1091.

[Hem82a] P.W. Hemker, *An accurate methods without directional bias for the numerical solution of a 2-d elliptic perturbation problem, in: Thy. and Appls. of Sing. Perts*, Springer LNM vol. 942 (eds.: W. Eckhaus and E.M Jaeger), Springer, Berlin, 1982.

[Hem82b] P.W. Hemker, *Mixed defect correction iteration for the accurate solution of the convection diffusion equations, in: Multigrid Methods*, Springer LNM vol. 960 (eds.: W. Hackbusch and V. Trottenberg), Springer, Berlin, 1982.

[HK88] P.W. Hemker and B. Koren, *Multigrid defect correction and upwind schemes for the steady Navier-Stokes equations, in Num. Meth. for Fluid Dynamics III*, Clarendon Press (eds.: K.W. Morton and M.J. Baines), Oxford, 1988.

[HK92] P.W. Hemker and B. Koren, *Defect correction and nonlinear multigrid for the steady Euler equations, in: Advances in C. F. D.*, Cambridge Univ. Press (eds.: W.G. Habashi and M.M. Hafez), Cambridge, 1992.

[Hou06] A. Householder, *The theory of matrices in numerical analysis*, Dover publications, 2006.

[J99] B. Juncu, *A numerical study of steady viscous flow past a fluid sphere*, Int. J. Heat Fluid Fl. 20(4) (1999), 414-421.

[KC79] J.T. King and D. Chillingworth, *Approximation of generalized inverses by iterated regularization,* Numer. Functional Anal. and Optim. 1(1979) 499-513.

[Koe84] J.J. Koenderink, *The structure of images*, *Biol. Cybernetics*, 50 (1984) 363-370.

[LL05] W. Layton and R. Lewandowski, *Residual stress of approximate deconvolution large eddy simulation models of turbulence. Journal of Turbulence*, 46(2): 1-21, 2006.

[L04] H.K. Lee, *Analysis of a defect correction method for viscoelastic fluid flow*, Comp. Math. Appls. 48(2004) 1213-1229..

[Lin94] T. Lindeberg, *Scale-space theory in computer vision, Kluwer, Dordrecht, 1994.*

[L94] J. Liu, *A sensitivity analysis for least squares ill-posed problems using the Haar basis*, SINUM 31(1994) 1486-1496.

[LR08] B. Lewis and L. Reichel, *Arnoldi-Tikhonov regularization methods*, J. Comput. Appl. Math., to appear, 2008.

[MP06] P. Mathe and S. V. Pereverzev, *Regularization of some noisy ill-posed problems with noisy, random data*, Math. Comp. 75(2006) 1913-1929.

[MN02] V. Nefedov and R.M.M. Mattheij, *Local defect correction with different grid types*, Numer. Meth. Part. Dif. Equ. 18(4) (2002), 454-468.

[N84] F. Natterer, *Error bounds for Tikhonov regularization in Hilbert scales*, Appl. Anal.

18(1984) 262-270.

[S06] E. R. Scheinerman, *C++ for Mathematicians: An Introduction for Students and Professionals*, CRC Press, 2006

[S07] T. Schuster, *The Method of Approximate Inverses*, Springer LNM 1906, Springer, Berlin, 2007.

[S05] T. Schuster, *Defect correction in vector field tomography: detecting the polynomial part using BEM and implementation of the method*, Inverse Problems 21 (2005) 75-91.

[SS02] L G Stanley and D L Stewart, *Design sensitivity analysis: Computational issues of sensitivity equation methods*, vol 25 in: Frontiers in Applied Math, SIAM, Philadelphia, 12002.

[S78] H.J. Stetter, *The defect correction principle and discretization methods*, Numer. Math. 29(1978) 425-443.

[V82] G. Vainikko, *Solution methods for linear ill-posed problems in Hilbert spaces*, Nauka, Moscow, 1982.

[VV86] G. Vainikko and A.Y. Veretennikov, *Iterative procedures in ill-posed problems*, Nauka, Moscow, 1986.

[V06] V.V. Vasin, *Some tendencies in Tikhonov regularization of ill-posed problems*, J. Inverse and Ill-posed Problems, 14(2006) 813-840.

[V02] C.R. Vogel, *Computational Methods for Inverse Problems*, SIAM publications, Philadelphia, 2002.