Georg-August-Universität
Göttingen

An utilization of a rough approximation of a noise covariance within the framework of multi-parameter regularization

Frank Bauer, Sergei Pereverzev

Preprint Nr. 2005-38

Preprint-Serie des
Instituts für Numerische und Angewandte Mathematik
Lotzestr. 16-18
D - 37083 Göttingen

An utilization of a rough approximation of a noise covariance within the framework of multi-parameter regularization

Frank Bauer[1)] and Sergei Pereverzev[2]

[1)] Institute for Numerical and Applied Mathematics
University of Göttingen, Department of Mathematics
Lotzestr. 16-18
37083 Göttingen
Germany
Email: bauer@mathematik.uni-kl.de,

[2)] Johann Radon Institute for Computational and Applied Mathematics (RICAM)
Austrian Academy of Sciences
Altenbergerstrae 69
A-4040 Linz
Austria
Email: sergei.pereverzyev@oeaw.ac.at

Abstract

Regularization under the assumption of badly known noise covariance operators is a demanding subject. In order to increase stability we propose a multi parameter regularization scheme and a parameter choice rule accordingly. We can show that the new scheme fulfills the same error bounds as the classically known ones.

Numerically we see that the new scheme is considerably more stable towards misestimations of the covariance operator than the old one-parameter one.

AMS-Classification: Primary 62G05; Secondary 62G20, 65J20

Keywords: Multi-Parameter Regularization, Balancing Principle, Optimal Rates

1 Introduction

In this article we will study ill-posed problems. We want to recover an element x_0 from some real Hilbert space $\mathcal{X}$ from indirectly observed data near $y_0 = Ax_0$ where A is an injective compact linear operator with $A : \mathcal{X} \to \mathcal{Y}$. The space $\mathcal{Y}$ is again a real Hilbert space; both are equipped with inner products $\langle \cdot, \cdot \rangle$ and norms $|| \cdot ||$ respectively.

Such inverse problems often arise in scientific context. One possible example is the downward continuation problem [FGS98] arising in satellite geodesy where x_0 is a gravitational potential on the Earth's surface and y_0 is the same potential observed on some satellite orbit.

In practice indirect observations are measured usually in the presence of some noise, so that we get y_ξ given by

$$y_\xi = Ax_0 + \xi, \tag{1}$$

where ξ is assumed to be a zero mean random element. This specifically means that for every element $f \in \mathcal{Y}$ we can observe $y_\xi(f) = \langle Ax_0, f \rangle + \xi(f)$ where $\xi(f) = \langle \xi, f \rangle$ is a random variable in a probability space $\{\Omega, \mathbb{P}\}$ with zero mean and variance $\left\langle K_\xi^2 f, f \right\rangle$. In this context an operator K_ξ^2 is called covariance operator and can be seen as a bounded self-adjoint non-negative operator from $\mathcal{Y} \to \mathcal{Y}$ sucht that for any $f, g \in \mathcal{Y}$ it holds $\mathbb{E} \langle \xi, f \rangle \langle \xi, g \rangle = \left\langle K_\xi^2 f, g \right\rangle = \langle K_\xi f, K_\xi g \rangle$ where $\mathbb{E}$ is the expectation with respect to the probability $\mathbb{P}$.

The observation equation (1) with random noise ξ means in fact that y_ξ cannot be observed exactly, but it can only be observed in a discretized form. To be more precise, instead of y_ξ we only have a vector $\left(y_\xi^1, y^2, \ldots, y_\xi^m\right) \in \mathbb{R}^m$ with coordinates

$$y_\xi^i = \langle y_\xi, f_i \rangle = \langle Ax_0, f_i \rangle + \langle \xi, f_i \rangle, \qquad i = 1, 2, \ldots, m, \tag{2}$$

determined by a design system $\{f_i\} \subset \mathcal{Y}$.

When fixing a design $\{f_i\}$ we may rewrite (2) as

$$P_m y_\xi = P_m A x_0 + P_m \xi, \tag{3}$$

where P_m denotes the orthogonal projection onto $\text{span}\{f_i\}_{i=1,\ldots,m}$. Note that (3) can be seen as a perturbed version of a discretized equation

$$P_m A x_0 = P_m y_0. \tag{4}$$

A minimal norm solution x_0^m of (4) has the property

$$||x_0 - x_0^m|| = \inf_{x \in \text{span}\{A^* f_i\}_{i=1,\ldots,m}} ||x_0 - x||. \tag{5}$$

Let

$$B_m = \sum_{i=1}^{m} s_i(B_m) \langle \varphi_i^m, \cdot \rangle \psi_i^m \tag{6}$$

be a singular value decomposition of the operator $B_m = P_m A$. Here $s_i(B_m)$, φ_i^m and ψ_i^m are such that

$$||P_m A|| = s_1(B_m) \geq s_2(B_m) \geq \cdots \geq s_m(B_m) > 0,$$

and

$$P_m A A^* P_m \psi_i^m = s_i^2(B_m)\psi_i^m, \quad A^* P_m A \varphi_i^m = s_i^2(B_m)\varphi_i^m.$$

Moreover $\{\psi_i^m\}$, $\{\varphi_i^m\}$ are the orthonormal bases of $\text{span}\{f_i\}_{i=1,\ldots,m}$ and $\text{span}\{A^* f_i\}_{i=1,\ldots,m}$ respectively. Then

$$x_0^m = \sum_{i=1}^{m} s_i^{-1}(B_m) \langle \psi_i^m, y_0 \rangle \varphi_i^m. \tag{7}$$

When dealing with discretized noisy observations (2) and (3) one can characterize a noise level $||P_m \xi||$ in terms of the trace of $P_m K_\xi^2 P_m$ which is a covariance operator for $P_m \xi$. Indeed

$$\mathbb{E}||P_m \xi||^2 = \sum_{i=1}^{m} \mathbb{E} \langle \psi_i^m, \xi \rangle^2 = \sum_{i=1}^{m} \left\langle K_\xi^2 \psi_i^m, \psi_i^m \right\rangle = \text{tr}(P_m K_\xi^2 P_m).$$

Then for $\varepsilon = \kappa \left(\text{tr}(P_m K_\xi^2 P_m) \right)^{1/2}$ a well-known Chebyshev inequality yields

$$\mathbb{P}\{||P_m \xi|| > \varepsilon\} = \mathbb{P}\{||P_m \xi||^2 > \varepsilon^2\} < \frac{\mathbb{E}||P_m \xi||^2}{\varepsilon^2} = \kappa^{-2}.$$

For example, with 95% confidence

$$||P_m \xi|| < 4.55 \left(\text{tr}(P_m K_\xi^2 P_m) \right)^{1/2}, \tag{8}$$

where we assume without loss of generality that $\text{tr}(P_m K_\xi^2 P_m) < 1$. Thus to estimate the noise level in (3) one just needs to know the trace of $P_m K_\xi^2 P_m$.

At the same time, in practice a covariance operator $P_m K_\xi^2 P_m$ is usually used for whitening or decorrelation of the noisy observations (3) (see e.g. [PP02]). After whitening an estimation of x_0 can be found by solving a regularized equation

$$\alpha x + A^* (P_m K_\xi^2 P_m)^{-1} A x = A^* (P_m K_\xi^2 P_m)^{-1} y_\xi, \tag{9}$$

which is a simple form of Tikhonov regularization with the regularization parameter α.

However, in practice one cannot expect that a covariance operator $P_m K_\xi^2 P_m$ will be known exactly. For example in satellite geodesy, as it has been indicated in [KK02], the measurement equipment on board of the satellite has never been validated in orbit. Furthermore measurements are usually contaminated with an aliasing signal caused by the unmodelled high frequencies.

As a result only some approximation $P_m \widetilde{K}_\xi^2 P_m$ of a covariance may be available which still allows to obtain reasonable estimates for $||P_m \xi||$, but its use instead of $P_m K_\xi^2 P_m$ in (9) leads to a very poor performance of the Tikhonov regularization as numerical experiments in [PP02] show.

The objective of the present paper is to discuss a new way of the use of an approximation $P_m\widetilde{K}_\xi^2 P_m$ of the covariance operator which does not involve it in the estimation process, but provides an almost optimal recovery of x_0 with a high confidence provided that the traces of $P_m K_\xi^2 P_m$ and $P_m\widetilde{K}_\xi^2 P_m$ are rather similar.

2 Multi-parameter Tikhonov regularization scheme

Using the orthonormal basis $\{\psi_i^m\}$ from (6) one can represent P_m as

$$P_m = \sum_{i=1}^{m} \psi_i^m \langle \psi_i^m, \cdot \rangle .$$

Moreover, for any finite sequence $\{m_n\}_{n=0,\dots M}$ of integers such that

$$0 = m_0 < m_1 < \cdots < m_M < m_{M+1} = m \tag{10}$$

the following orthogonal projections can be introduced ($n \in \{0, \dots, M\}$)

$$Q_n = \sum_{i=m_n+1}^{m_{n+1}} \psi_i^m \langle \psi_i^m, \cdot \rangle ,$$

which immediately yield

$$P_m = \sum_{n=0}^{M} Q_n. \tag{11}$$

We will call the traces $P_m K_\xi^2 P_m$ and $P_m\widetilde{K}_\xi^2 P_m$ similar if there are two constants $d_0, d_1 > 0$ and a finite sequence (10) such that for all $n \in \{0, \dots, M\}$ it holds

$$d_0^2 \operatorname{tr}(Q_n \widetilde{K}_\xi^2 Q_n) \le \operatorname{tr}(Q_n K_\xi^2 Q_n) \le d_1^2 \operatorname{tr}(Q_n \widetilde{K}_\xi^2 Q_n). \tag{12}$$

If $P_m K_\xi^2 P_m$ is a covariance operator of the random noise $P_m\xi$ and its trace is similar to the trace of $P_m\widetilde{K}_\xi^2 P_m$ then it holds

$$\mathbb{E}||Q_n\xi||^2 = \sum_{i=m_n+1}^{m_{n+1}} \mathbb{E}\langle \psi_i^m, \xi\rangle^2 = \sum_{i=m_n+1}^{m_{n+1}} \left\langle P_m K_\xi^2 P_m \psi_i^m, \psi_i^m \right\rangle = \operatorname{tr}(Q_n K_\xi^2 Q_n) \le d_1^2 \operatorname{tr}(Q_n \widetilde{K}_\xi^2 Q_n)$$

Returning to the 95% confidence region the norm $||Q_n\xi||$ can be estimated in terms of the trace of $P_m\widetilde{K}_\xi^2 P_m$ as

$$||Q_n\xi|| \le \delta_n, \qquad \delta_n = 4.55 d_1 \left(\sum_{i=m_n+1}^{m_{n+1}} \left\langle P_m K_\xi^2 P_m \psi_i^m, \psi_i^m \right\rangle \right)^{1/2} . \tag{13}$$

Moreover, with the same confidence we have

$$||P_m\xi|| \le \delta, \qquad \delta = \left(\sum_{n=0}^{M} \delta_n^2\right)^{1/2}. \tag{14}$$

Furthermore observe that the solution x_0^m of (4) admits a decomposition

$$x_0^m = \sum_{n=0}^{M} x_0^{m_n}, \tag{15}$$

where $x_0^{m_n}$ is a minimal norm solution of the equation

$$Q_n A x = Q_n y_0. \tag{16}$$

On the other hand it can be seen from (11) that the discretized observations (2) and (3) allow to construct a perturbed version of (16) for all $n \in \{0, \dots, M\}$

$$Q_n A x = Q_n y_\xi. \tag{17}$$

Then a Tikhonov regularization scheme with regularization parameter α_n applied to (17) leads to the equation

$$\alpha_n x + A^* Q_n A x = A^* Q_n y_\xi. \tag{18}$$

Its solution $x_{\alpha_n}^{m_n}$ can also be seen as the unique minimizer of the functional

$$||Q_n A x - Q_n y_\xi||^2 + \alpha_n ||S_n x||^2, \tag{19}$$

where

$$S_n = \sum_{i=m_n+1}^{m_{n+1}} \varphi_i^m \left\langle \varphi_i^m, \cdot \right\rangle,$$

and $\{\varphi_i^m\}$ is a system from (6) which is forming an orthonormal basis in $\mathrm{span}\{A^* f_i\}_{i=1,\dots,m} = \mathrm{span}\{A^* \psi_i^m\}_{i=1,\dots,m}$. It is easy to see that the vector

$$x_{\overrightarrow{\alpha}}^m = \sum_{n=0}^{M} x_{\alpha_n}^{m_n}, \qquad \overrightarrow{\alpha} = (\alpha_0, \dots, \alpha_M), \tag{20}$$

is the unique minimizer of the quadratic functional

$$J(\overrightarrow{\alpha}, x) = \sum_{n=0}^{M} \left(||Q_n A x - Q_n y_\xi||^2 + \alpha_n ||S_n x||^2\right) = ||P_m A x - P_m y_\xi||^2 + \sum_{n=0}^{M} \alpha_n ||S_n x||^2.$$

Following [BRZRS03] one can consider a minimization of $J(\overrightarrow{\alpha}, x)$ as a multi-parameter version of a standard Tikhonov regularization scheme. Several heuristically motivated recipes for the choice of $\overrightarrow{\alpha} = (\alpha_0, \dots, \alpha_M)$ have been discussed in [BRZRS03]. At the same time, in the considered case the estimates (13) and (14) together with the special structure of S_n and Q_n allow to choose α_n independently from each other. We will show that choosing these parameters in accordance with the strategy proposed recently in [MP03b] and [PS05] one can obtain an accuracy which is optimal up to some constant factor. One more advantage of our approach is related with a parallel treatment of the problems (17) that can be used for reducing computational complexity.

The regularized solution will approximate the true x_0^m quite well only if the regularization parameters are properly chosen dependent upon the noise level and the smoothness of x_0^m. Keeping in mind that x_0^m should fulfill the equation $B_m x = P_m y_0$ it is natural to describe its smoothness in terms of some smoothness index function Φ on the spectrum $B_m^* B_m$ by

$$x_0^m = \Phi(B_m^* B_m) v_0^m, \tag{21}$$

as it was suggested in [Heg92], [DES98] and [Tau98]. The function Φ here is continuous, increasing and satisfies $\Phi(0) = 0$ and v_0^m is an element from $\text{span}\{\varphi_i^m\}_{i=1,\dots,m}$.

For later use we will briefly summarize some results from the theory of Tikhonov regularization (see Th. 1, Th. 2, Prop. 3 in [MP03b])

Lemma 2.1
Let $\theta(t) = \Phi(t)\sqrt{t}$. Assume that the function $\Phi^2((\theta^2)^{-1}(t))$ is concave and the function $t \to t/\Phi(t)$ is non-decreasing for $t > 0$. If $z = \Phi(B^*B)v$, $||v|| \le \varrho$, and y_δ is such that $||Bz - y_\delta|| \le \delta$ then

$$||z - (\alpha I + B^* B)^{-1} B^* y_\delta|| \le \varrho \Phi(\alpha) + \frac{\delta}{2\sqrt{\alpha}}, \tag{22}$$

and for $\alpha_{opt} = \theta^{-1}(\frac{\delta}{2\rho})$ balancing the terms in the above estimate it holds

$$||z - (\alpha_{opt} I + B^* B)^{-1} B^* y_\delta||^2 \le 4\varrho^2 \Phi^2 \left((\theta^2)^{-1} \left(\frac{\delta^2}{4\varrho^2} \right) \right).$$

Remark
Note that the assumptions of lemma 2.1 are satisfied for index functions $\Phi(t) = t^\mu$, $0 < \mu \le 1$, respectively $\Phi(t) = \left(\ln t^{-1}\right)^{-\nu}$, $\nu > 0$, which are traditionally used in the regularization theory as smoothness index functions.

For the sequel we will assume that Φ fulfills the requirements of lemma 2.1 and consider a class $P_m A_\Phi(\rho)$ of solutions x_0^m which admit a representation (21) with $||v_0^m|| = \varrho$. If $x_0^m \in P_m A_\Phi(\rho)$ then from (6) it follows that

$$x_0^m = \sum_{i=1}^m \Phi(s_i^2(B_m)) \langle \varphi_i^m, v_0^m \rangle \varphi_i^m, \qquad \sum_{i=1}^m \langle \varphi_i^m, v_0^m \rangle^2 = \varrho^2. \tag{23}$$

Moreover (23) yields that the solution $x_0^{m_n}$ of (16) can be represented as

$$x_0^m = \sum_{i=m_n+1}^{m_{n+1}} \Phi(s_i^2(B_m)) \langle \varphi_i^m, v_0^m \rangle \varphi_i^m = \Phi(A^* Q_n A) v_0^{m_n}, \tag{24}$$

where

$$v_0^{m_n} = \sum_{i=m_n+1}^{m_{n+1}} \langle v_0^m, \varphi_i^m \rangle \varphi_i^m, \qquad ||v_0^{m_n}|| = \varrho_n, \qquad \sum_{n=0}^{M} \varrho_n^2 = \varrho^2. \tag{25}$$

In our notation it means that $x_0^{m_n} \in Q_n A_\Phi(\varrho_n)$.

Observe that if one applies a standard one parameter Tikhonov regularization to the noisy equation (3) then from lemma 2.1 it follows that

$$\sup_{x_0^m \in P_m A_\Phi(\varrho)} \sup_{\xi:\, ||P_m \xi|| \le \delta} ||x_0^m - (\alpha_{opt} I + B_m^* B_m)^{-1} B_m^* y_\xi||^2 \le 4\varrho^2 \Phi^2 \left((\theta^2)^{-1} \left(\frac{\delta^2}{4\varrho^2} \right) \right). \tag{26}$$

Similarly from lemma 2.1 and (17), (24) and (25) with $\alpha_n^{opt} = \theta^{-1}\left(\frac{\delta_n}{2\varrho_n}\right)$ we have

$$\sup_{x_0^{m_n} \in Q_n A_\Phi(\varrho_n)} \sup_{\xi:\, ||Q_n \xi|| \le \delta_n} ||x_0^{m_n} - x_{\alpha_n^{opt}}^{m_n}||^2 \le 4\varrho_n^2 \Phi^2 \left((\theta^2)^{-1} \left(\frac{\delta_n^2}{4\varrho_n^2} \right) \right). \tag{27}$$

Then with probability $(0.95)^M$ the accuracy of the regularized solution $x_{\vec{\alpha}_{opt}}^m$, where $\vec{\alpha}_{opt} = (\alpha_0^{opt}, \alpha_1^{opt}, \ldots \alpha_M^{opt})$ given by a multi parameter Tikhonov regularization can be estimated as

$$\begin{aligned}
\sup_{x_0^m \in P_m A_\Phi(\varrho)} \sup_{\xi:\, ||P_m \xi|| \le \delta} ||x_0^m - x_{\vec{\alpha}_{opt}}^m||^2 &= \sum_{n=0}^{M} \sup_{x_0^{m_n} \in Q_n A_\Phi(\varrho_n)} \sup_{\xi:\, ||Q_n \xi|| \le \delta_n} ||x_0^{m_n} - x_{\alpha_n^{opt}}^{m_n}||^2 \\
&\le \varrho^2 \sum_{n=0}^{M} 4 \frac{\varrho_n^2}{\varrho^2} \Phi^2 \left((\theta^2)^{-1} \left(\frac{\delta_n^2}{4\varrho_n^2} \right) \right) \\
&\le \varrho^2 4 \Phi^2 \left((\theta^2)^{-1} \left(\sum_{n=0}^{M} \frac{\delta_n^2 \varrho_n^2}{4\varrho_n^2 \varrho^2} \right) \right) \\
&= \varrho^2 4 \Phi^2 \left((\theta^2)^{-1} \left(\frac{\delta^2}{4\varrho^2} \right) \right),
\end{aligned} \tag{28}$$

here we use the concavity of $\Phi^2((\theta^2)^{-1}(t))$, (14) and (25).

Comparing (26) and (28) one can conclude that if $\Phi^2((\theta^2)^{-1}(t))$ is strictly concave then a multi-parameter Tikhonov regularization can potentially guarantee a better accuracy of estimation than the standard one-parameter one.

Remark

At first glance the last result means that the number of blocks M must not be too big. However in practice this will not pose too much trouble, we think. First of all, even if we are having problems in one block it is rather contained there and does not spread to the others and is comparably small because it is just a finite dimensional problem, i.e. although the likelihood that something is going wrong increases with the number of blocks, the damage which is imposed is decreasing.

With some much more involved treatment along the lines of [BP05] it seems to be possible to get similar results in expectation instead in probability. However comparing the multi-parameter Tikhonov with the one-parameter version in this setting would be out of the scope of this article.

Unfortunately an optimal a priori parameter choice $\overrightarrow{\alpha}_{opt} = \left(\theta^{-1}\left(\frac{\delta_n}{2\varrho_n}\right)\right)_{n=0,\dots,M}$ can seldomly used because the smoothness index function Φ is generally unknown.

Therefore, the question is how to choose the regularization parameter without the knowledge of Φ. An answer can be found in [PS05] where a general parameter choice strategy has been proposed. To describe this strategy we would like to remind that in practical applications the values of the regularization parameters are often selected from some finite geometric sequence, say

$$\Delta_T^n = \{\alpha_{n,i} = \delta_n^2 q^i, i = 0, 1, \dots, T\}, \qquad q > 1,$$

and corresponding regularized solutions $x_{\alpha_{n,i}}^{m_n}$ are studied online. In the considered case the strategy from [PS05] consists in the choice of $\alpha_n^+ \in \Delta_T^n$ as

$$\alpha_n^+ = \max\{\alpha_{n,i} \in \Delta_T^n \;:\; ||x_{\alpha_{n,i}}^{m_n} - x_{\alpha_{n,j}}^{m_n}|| \leq \frac{2\delta_n}{\sqrt{\alpha_{n,j}}},\ j = 0, \dots, i\}$$

Remark

In practice it might be sensible to use instead of $\frac{2\delta_n}{\sqrt{\alpha}}$ *the quantity* $2\kappa\,\mathrm{tr}(T_\alpha^n Q_n \widetilde{K}_\xi^2 Q_n (T_\alpha^n)^*)$ *where* $T_\alpha^n = (A^* Q_n A - \alpha S_n)^{-1} A$ *is the Tikhonov operator and* κ *is some tuning parameter (see e.g. [BP05]).*

The following statement is a direct consequence of Theorem 2.1 from [PS05]

Theorem 2.2

Let (12), (21) and the assumptions of lemma 2.1 be satisfied. Then with probability $(0.95)^M$ *the bounds (13), (14) for the noise levels hold for all* n *and*

$$||x_0^{m_n} - x_{\alpha_n^+}^{m_n}|| \leq 12 q \varrho_n \Phi(\alpha_n^{opt})$$

and in total

$$||x_0^m - x_{\overrightarrow{\alpha}_+}^m|| \leq 12q \left(\sum_{n=0}^{M} \varrho_n^2 \Phi^2(\alpha_n^{opt})\right)^{1/2},$$

where $\overrightarrow{\alpha}_+ = \left(\alpha_0^+, \alpha_1^+, \dots, \alpha_M^+\right)$.

Remark

Of course, the same adaptive parameter choice strategy can be applied to the one parameter Tikhonov regularization when

$$x_\alpha^m = \left(\alpha I + A^* P_m A\right)^{-1} A^* P_m y_\xi.$$

However, the same argument as in (28) shows that in this case the accuracy of the estimation cannot be better than for the multi-parameter regularization.

One more issue is also worth to be discussed. Recall that the original goal was to recover a solution x_0 of the equation $Ax = y_0$. So we are interested in estimation of the norm $||x_0 - x^m_{\vec{\alpha}_+}||$. This estimation can be derived from [MP03a] and [MP05] under the additional assumption that the smoothness index function $\Phi(t)$ is such that $\Phi^2(t)$ is operator monotone on the interval $[0, b]$ with $b > ||A||^2$.

Recall that the function Ψ is operator monotone on $[0, b]$ if for any pair of self-adjoint operators U and V with spectra in $[0, b]$ such that $U \leq V$ we also have $\Psi(U) \leq \Psi(V)$. As it has been observed in [MP05], if $x_0^m = \Phi(A^* P_m A) v_0^m$, $||v_0^m|| \leq \varrho$ and Φ^2 is operator monotone then $x_0^m = \Phi(A^*A)v_0$ with some $||v_0|| \leq \varrho$. Then in view of (5) it is natural to assume that x_0 can be also represented in the form $x_0 = \Phi(A^*A)w$, $||w|| \leq \varrho_1$. Moreover it is easy to see that the smoothness index function Φ is operator monotone if Φ^2 is.

Now using the formula (4) from [MP03a] with $g_\alpha(\lambda) = 1/\lambda$ and $\alpha = s_m^2(B_m)$ we obtain

$$||x_0 - x_0^m|| \leq c\left(\Phi(s_m(B_m)^2) + \Phi(||\,(I - P_m)\,A||^2)\right) \leq c_1 \Phi\left(||\,(I - P_m)\,A||^2\right),$$

where the constants c and c_1 do not depend on m. In particular we used the fact that

$$s_m(B_m) \leq s_m(A) \leq ||\,(I - P_m)\,A||.$$

Combining the estimation of $||x_0 - x_0^m||$ with theorem 2.2 we obtain

Theorem 2.3
Let the assumptions of lemma 2.1 and theorem 2.2 be satisfied. Assume that $\Phi(t)^2$ is operator monotone on $[0, b]$ with $b > ||A||^2$ and $x_0 = \Phi(A^*A)w$, $||w|| \leq \varrho_1$. Then with probability $(0.95)^M$

$$||x_0 - x^m_{\vec{\alpha}_+}|| \leq c\left(\Phi(||(I - P_m)A||^2) + \left(\sum_{n=0}^{M} \varrho_n^2 \Phi^2(\theta^{-1}(\frac{\delta_n}{2\varrho_n}))\right)^{1/2}\right),$$

where $\{\delta_n\}$ are the bounds for the noise levels (13) and c does neither depend on m nor on the $\{\delta_n\}$.

Before showing some numerical simulations we want to remark that in fact the multi-parameter Tikhonov is situated somewhere in between classical Tikhonov and Spectral-cutoff. Therefore we expect that at least in special situations the new method is not prone to saturation effects which would be another considerable advantage in comparison to standard Tikhonov.

3 Numerics

In order to test the described method we have set up the following example problem:

- A is a diagonal operator with 255 elements and Eigenvalue decay $k^{-4.5}$.
- Solution x_0 is a Gaussian random vector with Fourier coefficients decaying as k^{-2} along the Eigen spaces.
- Splitting is chosen along $m_n = 2^{n+1} - 1$ for $n = 1, \ldots, M$ where $M = 7$.

- The noise was chosen as a Gaussian random vector with variance $p^{-n} = 4^{-n}$ in the n-th Block, error level $\delta = 10^{-6}$
- For regularization we used the balancing principle; for more implementation details please have a look to either [BP05] or with a slightly advanced treatment a forthcoming article [BM05].
- The estimate of the covariance operator was chosen to be p^{-n} in the n-th block where we tried $p \in \{4.5, 5.0, 5.5\}$ which means we underestimate the error.
- We generate 50 different solutions and for each of them we also generate 10 different noisy input vectors. So we were treating 500 different problems for each case.
- As comparison method we chose the single parameter Tikhonov regularization with covariance correction and parameter choice performed as well by the balancing principle.

For each value of p we have displayed in the first bar diagram on the x-axis the logarithm of the relative error made for multi respectively single parameter Tikhonov; the y-axis contains the number of problems falling in the particular case. The second plot always shows the logarithmic ratio between the relative errors of multi- and single parameter Tikhonov.

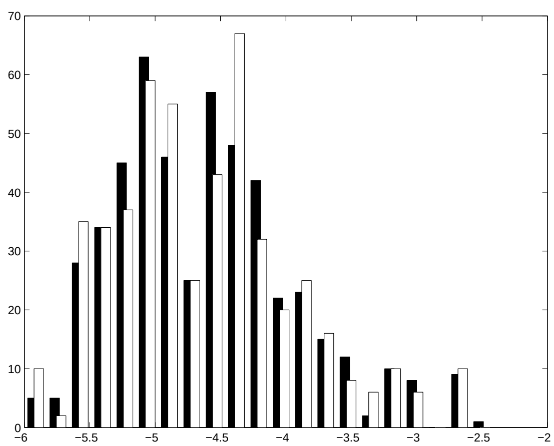

Figure 1: Case $p = 4.5$, log(relative error) (white: multi, black: single)

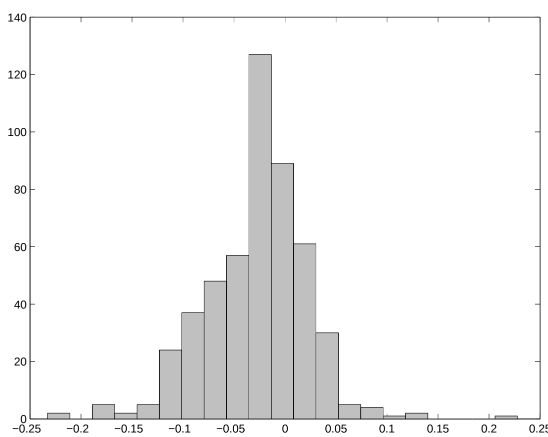

Figure 2: Case $p = 4.5$, log(relative error ratio multi/single)

3.1 Discussion

We observe the following facts:

- If we just slightly underestimated the error (Case $p = 4.5$) both methods roughly behave in the same way; sometimes one a bit better, sometimes the other a bit better but nothing significant

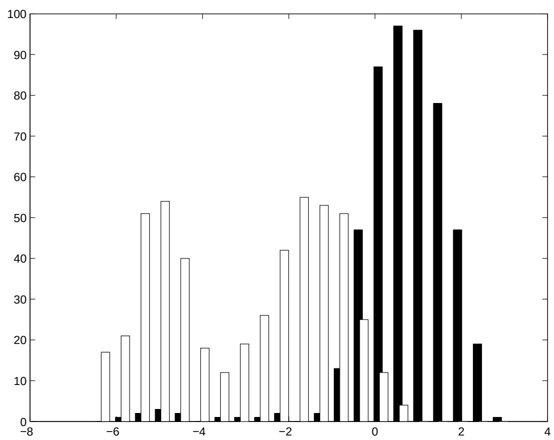

Figure 3: Case $p = 5.0$, log(relative error) (white: multi, black: single)

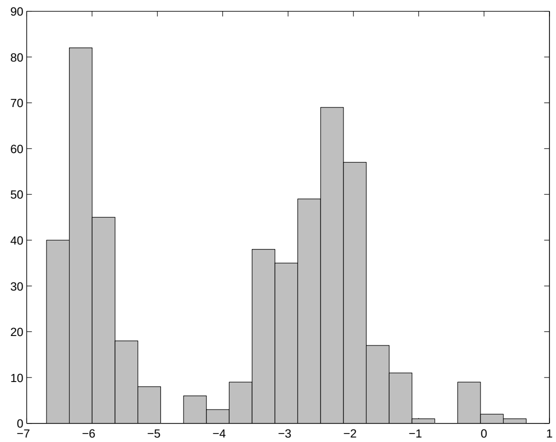

Figure 4: Case $p = 5.0$, log(relative error ratio multi/single)

- If we underestimate the error considerably (Case $p = 5.0$) the situation changes completely. The multi-parameter Tikhonov still works and the one parameter Tikhonov underregularizes the problem heavily. Therefore there is a difference of about 4 orders of magnitude between the results. The two bumps are most likely due to the fact that sometimes all blocks are regularized correctly and sometimes one block fails.

- When we underestimate the error heavily (Case $p = 5.5$) both methods do not work any more (as expected). However there is still a qualitative difference between the two methods. The multi-parameter Tikhonov is about two orders of magnitude better than the one parameter scheme.

Concluding we have that the newly proposed method works very well, even in expectation. As long as the estimation of the error is not too bad it is competitive two one of the mostly widespread used methods. When the error estimation gets worse there is a region where the new method still works whereas the old one-parameter scheme bails out. Thus, the new method seems to be much more robust in comparison to the old one.

Interestingly according to our numerical simulations it is likely that the difference between the solution of the new and the old method is a good indicator whether the estimate of the covariance is good or not. Every time when it was bad the difference was large, i.e. we had (statistically seen) no wrong pairs of solutions which were close together.

Acknowledgements

The first author wants to thank for the support of the Graduiertenkolleg 1023, University of Gttingen and the Radon Institute for Computational and Applied Mathematics for the invitation to the special semester on computational mechanics which made this article possible.

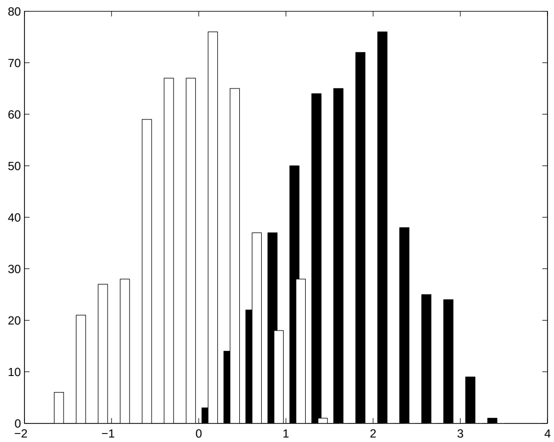

Figure 5: Case $p = 5.5$, log(relative error) (white: multi, black: single)

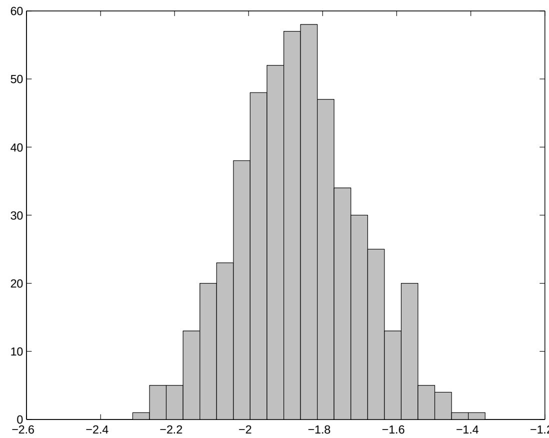

Figure 6: Case $p = 5.5$, log(relative error ratio multi/single)

References

[BM05] Bauer, F., and Munk, A., *Optimal Regularization for Moderately Ill-Posed Problems in Normed Spaces (forthcoming).*

[BP05] Bauer, F., and Pereverzev, S., *Regularization without Preliminary Knowledge of Smoothness and Error Behavior*, European Journal of Applied Mathematics **16** (2005), no. 3, 303–317.

[BRZRS03] Brezinski, C., Redivo-Zaglia, M., Rodriguez, G., and Seatzu, S., *Multi-parameter regularization techniques for ill-conditioned linear systems*, Numer. Math. **94** (2003), no. 2, 203–228.

[DES98] Deuflhard, P., Engl, H., and Scherzer, O., *A convergence analysis of iterative methods for the solution of nonlinear ill-posed problems under affinely invariant conditions*, Inverse Probl. **14** (1998), no. 5, 1081–1106.

[FGS98] Freeden, W., Gervens, T., and Schreiner, M., *Constructive Approximation on the Sphere (With Applications to Geomathematics)*, Numerical Mathematics and Scientific Computation, Oxford University Press, Oxford, New York, 1998.

[Heg92] Hegland, M., *An optimal order regularization method which does not use additional smoothness assumptions*, SIAM J. Numer. Anal. **29** (1992), no. 5, 1446–1461.

[KK02] Kusche, J., and Klees, R., *Regularization of gravity field estimation from satellite gravity gradients*, J. Geod. **76** (2002), 359–368.

[MP03a] Mathé, P., and Pereverzev, S., *Discretization strategy for linear ill-posed problems in variable Hilbert scales*, Inverse Problems **19** (2003), no. 6, 1263–1277.

[MP03b] Mathé, P., and Pereverzev, S., *Geometry of Linear Ill-Posed Problems in Variable Hilbert Spaces*, Inverse Problems **19** (2003), no. 3, 789–803.

[MP05] Mathé, P., and Pereverzev, S., *Regularization of some linear inverse problems with discretized random noise data (accepted)*, Math. Comput. (2005).

[PP02] Pail, R., and Plank, G., *Assessment of three numerical solution strategies for gravity field recovery from GOCE satellite gravity gradiometry implemented on a parallel platform.*, Journal of Geodesy **76** (2002), 462–474.

[PS05] Pereverzev, S., and Schock, E., *On the adaptive selection of the parameter in regularization of ill-posed problems (accepted)*, SIAM J. Numer. Anal. (2005).

[Tau98] Tautenhahn, U., *Optimality for ill-posed problems under general source conditions*, Numer. Funct. Anal. Optimization **19** (1998), no. 3–4, 377–398.

Institut für Numerische und Angewandte Mathematik
Universität Göttingen
Lotzestr. 16-18
D - 37083 Göttingen

Telefon: 0551/394512
Telefax: 0551/393944

Email: trapp@math.uni-goettingen.de URL: http://www.num.math.uni-goettingen.de

Verzeichnis der erschienenen Preprints:

2005-01	A. Schöbel, S. Scholl:	Line Planning with Minimal Traveling Time
2005-02	A. Schöbel	Integer programming approaches for solving the delay management problem
2005-03	A. Schöbel	Set covering problems with consecutive ones property
2005-04	S. Mecke, A. Schöbel, D. Wagner	Station location - Complex issues
2005-05	R. Schaback	Convergence Analysis of Methods for Solving General Equations
2005-06	M. Bozzini, L. Lenarduzzi, R. Schaback	Kernel B–Splines and Interpolation
2005-07	R. Schaback	Convergence of Unsymmetric Kernel–Based Meshless Collocation Methods
2005-08	T. Hohage, F.J. Sayas	Numerical solution of a heat diffusion problem by boundary element methods using the Laplace transform
2005-09	T. Hohage	An iterative method for ‘ inverse medium scattering problems based on factorization of the far field operator
2005-10	F. Bauer, T. Hohage	A Lepskij-type stopping rule for regularized Newton methods
2005-11	V. Dolean, F. Nataf, G. Rapin	New constructions of domain decomposition methods for systems of PDEs Nouvelles constructions de méthodes de décomposition de domaine pour des systèmes d'équations aux dérivées partielles
2005-12	F. Nataf, G. Rapin	Construction of a New Domain Decomposition Method for the Stokes Equations

2005-13	Y.C. Hon, R. Schaback	Solvability of Partial Differential Equations by Meshless Kernel Methods
2005-14	F. Bauer, P. Mathé, S. Pereverzev	Local Solutions to Inverse Problems in Geodesy: The Impact of the Noise Covariance Structure upon the Accuracy of Estimation
2005-15	O. Ivanyshyn, R. Kress	Nonlinear integral equations for solving inverse boundary value problems for inclusions and cracks
2005-16	R. Kress, W. Rundell	Nonlinear integral equations and the iterative solution for an inverse boundary value problem
2005-17	A. Yapar, H. Sahintürk, I. Akduman, R. Kress	One-Dimensional Profile Inversion of a Cylindrical Layer with Inhomogeneous Impedance Boundary: A Newton-Type Iterative Solution
2005-18	J. Puerto, A. Schöbel, S. Schwarze	The Path Player Game: Introduction and Equilibri
2005-19	R. Ahrens, A. Beckert, H. Wendland	A meshless spatial coupling scheme for large-scale fluid-structure-interaction problems
2005-20	F.J. Narcowich, J.D. Ward, X. Sun, H. Wendland	Direct and Inverse Sobolev Error Estimates for Scattered Data Interpolation via Spherical Basis Functions
2005-21	F.J. Narcowich, J.D. Ward, H. Wendland	Sobolev Error Estimates and a Bernstein Inequality for Scattered Data Interpolation via Radial Basis Functions
2005-22	I. Akduman, R. Kress, A. Yapar	Iterative Reconstruction of Dielectric Rough Surface Profiles at Fixed Frequency
2005-23	A. Schöbel, S. Scholl et.al.	DisKon - Disposition und Konfliktlsungsmanagement fr die beste Bahn
2005-24	P. Ivanyshyn, R. Kress	Nonlinear Integral Equations in Inverse Obstacle Scattering
2005-25	R. Kress, P. Serranho	A hybrid method for two-dimensional crack reconstruction
2005-26	R. Kress, P. Serranho	A hybrid method for sound-hard obstacle reconstruction
2005-27	G. Lube, G. Rapin	Residual-based stabilized higher-order FEM for advection-diffusion problems
2005-28	G. Lube, G. Rapin	Residual-based stabilized higher-order FEM for a generalized Oseen problem
2005-29	G. Lube, T. Knopp et.al.	Domain decomposition methods for indoor air flow simulation

2005-30	G. Lube et. al.	Stabilized finite element methods for the generalized Oseen problem
2005-31	R. Schaback	Multivariate interpolation by polynomials and radial basis functions
2005-32	R. Schaback	Inverse and Stability Estimates for Meshless Approximations by Translates of Kernel Functions
2005-33	R. Schaback	Kernel Techniques: From Machine Learning to Meshless Methods
2005-34	T. Hohage	Fast numerical solution of the electromagnetic medium scattering problem and applications to the inverse problem
2005-35	H. Harbrecht, T. Hohage	Fast Methods for Three-Dimensional Inverse Obstacle Scattering Problems
2005-36	R. Velásquez, M. Ehrgott, A. Schbel	A Set-packing Approach to Routing Trains Through Railway Station
2005-37	D. Colton, R. Kress	Using fundamental solutions in inverse scattering
2005-38	F. Bauer, S. Pereverzev	An utilization of a rough approximation of a noise covariance within the framework of multi-parameter regularization