Journal of Computational Mathematics, Vol.26, No.1, 2008, 37–55.

MULTI-PARAMETER TIKHONOV REGULARIZATION FOR LINEAR ILL-POSED OPERATOR EQUATIONS*

Zhongying Chen
Department of Scientific Computing and Computer Applications, Zhongshan University, Guangzhou 510275, China
Email: lnsczy@mail.sysu.edu.cn
Yao Lu and Yuesheng Xu
Department of Mathematics, Syracuse University, Syracuse, NY 13244-1150, USA
Email: ylu08@syr.edu, yxu06@syr.edu
Hongqi Yang
Department of Scientific Computing and Computer Applications, Zhongshan University, Guangzhou 510275, China
Email: mcsyhq@mail.sysu.edu.cn

Abstract

We consider solving linear ill-posed operator equations. Based on a multi-scale decomposition for the solution space, we propose a multi-parameter regularization for solving the equations. We establish weak and strong convergence theorems for the multi-parameter regularization solution. In particular, based on the eigenfunction decomposition, we develop a posteriori choice strategy for multi-parameters which gives a regularization solution with the optimal error bound. Several practical choices of multi-parameters are proposed. We also present numerical experiments to demonstrate the outperformance of the multi-parameter regularization over the single parameter regularization.

Mathematics subject classification: 47A52.
Key words: Ill-posed problems, Tikhonov regularization, Multi-parameter regularization.

1. Introduction

The classical regularization method for solving ill-posed problems, which was proposed independently by Phillips [18] and Tikhonov [22], has been proved to be an excellent idea to overcome the difficulty caused by the ill-posedness, see [9, 12, 19, 23]. This regularization method turns an ill-posed problem to a well-posed problem which can be efficiently solved by standard numerical methods (cf. [1]). Such a method using a *single* parameter regularization is based on the hypothesis that noise *effect* to an ill-posed problem is uniformly distributed in all frequency bands of the solution. The single parameter regularization method adds a uniform penalty to every frequency band of the solution or the high-frequency band of the solution. The first case may result in solutions that are too smooth to preserve certain features of the original data. In the second case, the regularization solutions may be affected by low-frequency noise.

In practice, we observe different circumstances which lead us to consider *multi-parameter* regularization. Often, noise distributes differently in different parts of the physical domain. There is a case when noise distributes differently in different frequency bands. Sometimes noise has different effects to different frequency bands (scales) of the solution even though the noise is uniformly distributed. These circumstances suggest an introduction of multi-parameters to

* Received November 16, 2006 / Revised version received February 28, 2007 / Accepted March 30, 2007 /

the regularization method. Multi-parameter regularization has been used in treating systems of linear equations in a few different contexts. A choice of multiple parameters was proposed in [2] by using the generalized L-curve method. A multi-parameter regularization algorithm for the solution of over-determined, ill-conditioned linear systems was proposed in [3], where numerical examples were presented to demonstrate that the proposed algorithm is stable and robust. In [8], the authors used a multi-parameter regularization method for atmospheric remote sensing. A multi-parameter regularization was used in [13] for solving a deconvolution problem in signal analysis when a wavelet transform was used to represent the system. The paper [14] proposed to use multi-parameter regularization methods based on biorthogonal wavelets and tight frame filter banks arising from the blurring kernel for treating the ill-posed problem related to high-resolution image reconstruction.

It is the main purpose of this paper to present convergence analysis for the multi-parameter regularization method for solving ill-posed operator equations when the solution space has a multi-scale decomposition (cf. [6, 10]). This paper will be based on the hypotheses that the function space and operators have a multi-scale structure and that noise has a different effect to a different frequency band of the solution due to the multi-scale structure of the solution, even though the noise is uniformly distributed. The proposed multi-parameter regularization will add different penalty parameters to different scales of the solution so that the ill-posedness is treated efficiently. At this point, we would like to point out that the multi-parameter regularization is more effective than single parameter regularization only if more information on the operator and the noise such as multi-scale decomposition is available.

The paper is organized into four sections. We describe in Section 2 the multi-parameter regularization method. Section 3 is devoted to the development of weak and strong convergence for the multi-parameter regularization solution. We also present an error estimate for the regularization solution and obtain a special result for regularization using the eigenfunction decomposition. In Section 4, we suggest a posteriori strategy for the choice of multi-parameters which gives a regularization solution with the optimal error bound when the eigenfunction decomposition is used. A numerical example is presented to illustrate the efficiency of this strategy. We also propose in Section 4 several practical strategies for the choice of the multiple parameters for the finite dimensional case, and present three numerical experiments in signal deconvolution and denoising using the multi-parameter regularization method. These experiments demonstrate the outperformance of the multi-parameter regularization over the single parameter regularization.

2. Multi-parameter Regularization Methods

We introduce in this section the multi-parameter regularization method for solving linear ill-posed operator equations based on a multi-scale decomposition of the solution space.

We first describe the linear ill-posed problem that we consider in this paper and recall the classical Tikhonov regularization method. Let $\mathbb{X}$ and $\mathbb{Y}$ be two Hilbert spaces. We will use $(\cdot,\cdot)$ for the inner product and $\|\cdot\|$ for the norm in both spaces without distinguishing them. Suppose that $\mathcal{K}:\mathbb{X}\to\mathbb{Y}$ is a linear compact operator. For a function $f\in\mathbb{Y}$, we consider the operator equation of the first kind

$$\mathcal{K}u = f. \tag{2.1}$$

We assume that the range $R(\mathcal{K})$ of operator $\mathcal{K}$ is of infinite dimension and thus, the solution of (2.1) does not continuously depend on the right-hand side f, that is, equation (2.1) is ill-posed

[12]. In this paper, without loss of generality we assume that $f \in R(\mathcal{K})$ and the operator $\mathcal{K}$ is injective, which means the nullspace $N(\mathcal{K}) = \{0\}$.

In practice, the exact data $f \in \mathbb{Y}$ may not be available. Instead, one may have a noisy data $f^\delta \in \mathbb{Y}$ with a known error level $\delta > 0$, i.e., $\|f^\delta - f\| \le \delta$. Actually, we have to solve the perturbed equation

$$\mathcal{K}u^\delta = f^\delta. \tag{2.2}$$

Since its solution does not continuously depend on the right-hand side data f^δ, regularization is necessary. The Tikhonov regularization method for equation (2.2) is to seek a stable approximate solution from the following equation for an appropriate positive parameter α

$$(\alpha\mathcal{I} + \mathcal{K}^*\mathcal{K})u_\alpha^\delta = \mathcal{K}^* f^\delta, \tag{2.3}$$

where $\mathcal{K}^*$ is the adjoint operator of $\mathcal{K}$. The parameter α is called the regularization parameter. This is the classical *single* parameter regularization which imposes a *uniform* penalty parameter. It works well when the operator $\mathcal{K}$ and the given data f have only *single* scale representations.

When *multi-scale* representations of the operator $\mathcal{K}$ and the given data f are available, one should make use of the multi-scale structure to impose different penalty parameters for different scales aiming at a better regularization. Such a method is called a *multi-parameter regularization* method, which we formulate precisely below. Let $\mathbb{N}_0 := \{0, 1, \cdots\}$ and for $\mathbb{A}, \mathbb{B} \subseteq \mathbb{X}$, let $\mathbb{A} \oplus^\perp \mathbb{B}$ denote the direct sum of subspaces $\mathbb{A}$ and $\mathbb{B}$ with $\mathbb{A} \perp \mathbb{B}$. Suppose that the space $\mathbb{X}$ has a multi-scale decomposition, that is,

$$\mathbb{X} := \bigoplus_{i\in\mathbb{N}_0}^{\perp} \mathbb{W}_i. \tag{2.4}$$

The spaces $\mathbb{W}_i$ could be generated by wavelet functions, by other orthogonal systems or by eigenfunctions of the operator $\mathcal{K}^*\mathcal{K}$. See [7] for a general reference on wavelet analysis and see [16, 17] for construction of wavelets on bounded domains which are particularly useful for solving integral equations.

Let $\mathcal{Q}_i$ denote the orthogonal projection from $\mathbb{X}$ onto $\mathbb{W}_i$, $i \in \mathbb{N}_0$. As a result, the identity operator $\mathcal{I}$ can be written as $\mathcal{I} := \sum_{i\in\mathbb{N}_0} \mathcal{Q}_i$. Using these notations, we have two ways to express a vector $v \in \mathbb{X}$ either as $v := \sum_{i\in\mathbb{N}_0} \mathcal{Q}_i v \in \mathbb{W}_0 \oplus^\perp \mathbb{W}_1 \oplus^\perp \cdots$, or $v := [\mathcal{Q}_0 v, \mathcal{Q}_1 v, \cdots]^T \in \mathbb{W}_0 \times \mathbb{W}_1 \times \cdots$. Following [4, 5], the operator $\mathcal{K}^*\mathcal{K} : \mathbb{X} \to \mathbb{X}$ can then be identified in the matrix notation as

$$\mathcal{K}^*\mathcal{K} = \begin{bmatrix} \mathcal{Q}_0\mathcal{K}^*\mathcal{K}\mathcal{Q}_0 & \mathcal{Q}_0\mathcal{K}^*\mathcal{K}\mathcal{Q}_1 & \cdots & \mathcal{Q}_0\mathcal{K}^*\mathcal{K}\mathcal{Q}_k & \cdots \\ \mathcal{Q}_1\mathcal{K}^*\mathcal{K}\mathcal{Q}_0 & \mathcal{Q}_1\mathcal{K}^*\mathcal{K}\mathcal{Q}_1 & \cdots & \mathcal{Q}_1\mathcal{K}^*\mathcal{K}\mathcal{Q}_k & \cdots \\ \vdots & \vdots & \ddots & \vdots & \\ \mathcal{Q}_k\mathcal{K}^*\mathcal{K}\mathcal{Q}_0 & \mathcal{Q}_k\mathcal{K}^*\mathcal{K}\mathcal{Q}_1 & \cdots & \mathcal{Q}_k\mathcal{K}^*\mathcal{K}\mathcal{Q}_k & \cdots \\ \vdots & \vdots & & \vdots & \ddots \end{bmatrix}. \tag{2.5}$$

For nonnegative numbers $\lambda_i, i \in \mathbb{N}_0$, let Λ be an operator from $\mathbb{X}$ to $\mathbb{X}$ defined by

$$\Lambda := \sum_{i\in\mathbb{N}_0} \lambda_i \mathcal{Q}_i, \tag{2.6}$$

which has the matrix form

$$\Lambda = \begin{bmatrix} \lambda_0\mathcal{Q}_0 & 0 & \cdots \\ 0 & \lambda_1\mathcal{Q}_1 & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix}. \tag{2.7}$$

We call equations

$$(\Lambda + \mathcal{K}^*\mathcal{K})u_\Lambda = \mathcal{K}^* f \tag{2.8}$$

and

$$(\Lambda + \mathcal{K}^*\mathcal{K})u_\Lambda^\delta = \mathcal{K}^* f^\delta \tag{2.9}$$

the multi-parameter regularization method for equations (2.1) and (2.2), respectively, and call $[\lambda_i : i \in \mathbb{N}_0]^T$ the regularization parameter vector. This method allows us to choose different parameters according to the multi-scale behavior of the operators and the perturbations of the data in different scales. Here we consider an infinitely dimensional regularization parameter vector for theoretical interest in order to better understand the insight of this method, while, in practice, only finite dimensional regularization parameter vector will be used. We will discuss this later in the next section. Introducing $\lambda_- := \inf\{\lambda_i : i \in \mathbb{N}_0\}$ and $\lambda_+ := \sup\{\lambda_i : i \in \mathbb{N}_0\}$, when $\lambda_- = \lambda_+ = \alpha > 0$, the multi-parameter regularization scheme (2.9) reduces to the single parameter regularization scheme (2.3).

We say that Λ has property (**A**) if there exists a positive constant c_Λ such that for all $x \in \mathbb{X}$

$$((\Lambda + \mathcal{K}^*\mathcal{K})x, x) \geq c_\Lambda \|x\|^2. \tag{2.10}$$

If Λ has property (**A**), then the inverse operator $(\Lambda + \mathcal{K}^*\mathcal{K})^{-1}$ exists and has the estimate

$$\|(\Lambda + \mathcal{K}^*\mathcal{K})^{-1}\| \leq 1/c_\Lambda. \tag{2.11}$$

We remark that if $\lambda_- > 0$, then Λ has property (**A**) with $c_\Lambda := \lambda_-$. Moreover, for any $f \in \mathbb{Y}$, (2.8) has unique solution $u_\Lambda := \mathcal{R}_\Lambda f$, which depends continuously on f, where $\mathcal{R}_\Lambda := (\Lambda + \mathcal{K}^*\mathcal{K})^{-1}\mathcal{K}^*$. It can be proved that for any $f \in \mathbb{Y}$, $u_\Lambda \in \mathbb{X}$ is the solution of (2.8) if and only if u_Λ is the minimizer of the functional

$$F(u) := \|\mathcal{K}u - f\|^2 + \sum_{i \in \mathbb{N}_0} \lambda_i \|\mathcal{Q}_i u\|^2, \quad u \in \mathbb{X}. \tag{2.12}$$

Note that if λ_i is chosen as zero it means that we do not impose a regularization penalty in the scale corresponding to i and if λ_i is chosen as $+\infty$ it means $\mathcal{Q}_i u = 0$. In the latter case, the component of the regularization solution corresponding to scale i is equal to zero.

3. Convergence Analysis

In this section, we present convergence results and error estimates for the regularization methods (2.8) and (2.9) for solving equation (2.1). The first result is a weak convergence of the regularization method (2.8) with the infinite number of regularization parameters λ_i. This result also prepares us for proving the main result of this section, the strong convergence of the regularization method (2.8) with a *finite* number of regularization parameters Λ. As we pointed out in Section 1 and as we will demonstrate in numerical experiments, the case of using only finite number of parameters is the most important one in applications. When the regularization operator Λ is a self-adjoint bounded and positive semi-definite linear operator not necessarily in the form (2.7) we establish an error estimate of the regularization solution if Λ commutes with $\mathcal{K}$. We also present error expressions of the regularization solution when the multi-scale decomposition of space $\mathbb{X}$ is given by the eigenfunctions of $\mathcal{K}^*\mathcal{K}$.

We now present a weak convergence theorem.

Theorem 3.1. *Suppose that* $\mathcal{K}:\mathbb{X}\to\mathbb{Y}$ *is a linear injective compact operator and* $\Lambda:\mathbb{X}\to\mathbb{X}$ *is defined by* (2.6) *having property (**A**). Let* u_* *be the solution of* (2.1) *and* u_Λ *the solution of* (2.8). *If* $\lambda_+/c_\Lambda=\mathcal{O}(1)$ *as* $\lambda_+\to 0$*, then* u_Λ *converges weakly to* u_**, that is,* $u_\Lambda \rightharpoonup u_*$*, as* $\lambda_+\to 0$*, in the sense that* $(u_\Lambda,x)\to(u_*,x)$*, as* $\lambda_+\to 0$*, for all* $x\in\mathbb{X}$*.*

Proof. We denote by e_Λ the error of the regularization solution, that is, $e_\Lambda := u_* - u_\Lambda$. It follows from property (**A**) that the inverse $(\Lambda+\mathcal{K}^*\mathcal{K})^{-1}$ exits and thus by (2.8) we have that

$$e_\Lambda = u_* - (\Lambda+\mathcal{K}^*\mathcal{K})^{-1}\mathcal{K}^*f = (\Lambda+\mathcal{K}^*\mathcal{K})^{-1}\Lambda u_*. \tag{3.1}$$

Therefore, by (2.11) and the hypotheses of this theorem, there exists a positive constant c independent of Λ such that

$$\|e_\Lambda\| = \|(\Lambda+\mathcal{K}^*\mathcal{K})^{-1}\Lambda u_*\| \le \lambda_+/c_\Lambda\|u_*\| \le c\|u_*\|. \tag{3.2}$$

This ensures that there exist a subsequence Λ_j of Λ converging to 0 and some $v\in\mathbb{X}$ such that e_{Λ_j} converges weakly to v, that is,

$$e_{\Lambda_j} \rightharpoonup v, \quad \text{as} \quad j\to\infty. \tag{3.3}$$

Since for each positive integers j, $(\Lambda_j+\mathcal{K}^*\mathcal{K})e_{\Lambda_j} = \Lambda_j u_*$, we have that

$$((\Lambda_j+\mathcal{K}^*\mathcal{K})e_{\Lambda_j}, e_{\Lambda_j}) = (\Lambda_j u_*, e_{\Lambda_j}), \tag{3.4}$$

which together with (3.2) and the fact that $(\Lambda_j e_{\Lambda_j}, e_{\Lambda_j}) \ge 0$ yields

$$\|\mathcal{K}e_{\Lambda_j}\|^2 \le (\Lambda_j u_*, e_{\Lambda_j}) \le c\|u_*\|^2\|\Lambda_j\|. \tag{3.5}$$

Letting $j\to\infty$ in (3.5) we find that

$$\lim_{j\to\infty}\|\mathcal{K}e_{\Lambda_j}\| = 0. \tag{3.6}$$

Noticing that a compact linear operator maps a weak convergence sequence to a strong convergence sequence, since $\mathcal{K}$ is a compact linear operator, it follows from (3.3) that

$$\lim_{j\to\infty}\|\mathcal{K}e_{\Lambda_j} - \mathcal{K}v\| = 0.$$

This with (3.6) yields

$$\mathcal{K}v = 0. \tag{3.7}$$

Because of the injectivity of $\mathcal{K}$ we conclude from (3.7) that $v=0$. Thus, it follows from (3.3) that $e_{\Lambda_j} \rightharpoonup 0$, as $j\to\infty$.

We next show by contradiction that we in fact have the result

$$e_\Lambda \rightharpoonup 0, \quad \text{as} \quad \lambda_+\to 0. \tag{3.8}$$

If equation (3.8) does not hold, there must be a subsequence e_{Λ_j} of e_Λ which converges weakly to a limit different from zero as Λ_j converges to 0. This subsequence is also bounded and thus, by repeating the proof above, we can prove that itself has a subsequence that converges weakly to zero. This is a contradiction and completes the proof of this theorem.

We now consider the regularization operator having only finite number of parameters. In practice, using only finite number of parameters is more interesting. For example, applications considered in [13, 14] use only finite number of parameters. Specifically, for a fixed positive integer N we assume that the space $\mathbb{X}$ has the decomposition

$$\mathbb{X} = \bigoplus_{i\in\mathbb{Z}_{N+1}}^{\perp} \mathbb{W}_i, \tag{3.9}$$

where $\mathbb{Z}_{N+1} := \{0, 1, \cdots, N\}$. In this setting, there are two possibilities. The space $\mathbb{X}$ is a finite dimensional space or the space $\mathbb{W}_N$ may be of infinite dimension. In the second case, the space $\mathbb{W}_N$ combines all subspaces $\mathbb{W}_i$ in (2.4) with indices $i \geq N$. Corresponding to the decomposition (3.9), operators $\mathcal{K}^*\mathcal{K}$ and Λ become

$$\mathcal{K}^*\mathcal{K} = \begin{bmatrix} \mathcal{Q}_0\mathcal{K}^*\mathcal{K}\mathcal{Q}_0 & \mathcal{Q}_0\mathcal{K}^*\mathcal{K}\mathcal{Q}_1 & \cdots & \mathcal{Q}_0\mathcal{K}^*\mathcal{K}\mathcal{Q}_N \\ \mathcal{Q}_1\mathcal{K}^*\mathcal{K}\mathcal{Q}_0 & \mathcal{Q}_1\mathcal{K}^*\mathcal{K}\mathcal{Q}_1 & \cdots & \mathcal{Q}_1\mathcal{K}^*\mathcal{K}\mathcal{Q}_N \\ \vdots & \vdots & \ddots & \vdots \\ \mathcal{Q}_N\mathcal{K}^*\mathcal{K}\mathcal{Q}_0 & \mathcal{Q}_N\mathcal{K}^*\mathcal{K}\mathcal{Q}_1 & \cdots & \mathcal{Q}_N\mathcal{K}^*\mathcal{K}\mathcal{Q}_N \end{bmatrix}, \tag{3.10}$$

and

$$\Lambda = \sum_{i\in\mathbb{Z}_{N+1}} \lambda_i \mathcal{Q}_i = \begin{bmatrix} \lambda_0\mathcal{Q}_0 & 0 & \cdots & 0 \\ 0 & \lambda_1\mathcal{Q}_1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_N\mathcal{Q}_N. \end{bmatrix}, \tag{3.11}$$

respectively. In this setting, we prove the next strong convergence result.

Theorem 3.2. *Suppose that $\mathcal{K} : \mathbb{X} \to \mathbb{Y}$ is a linear injective compact operator, and $\Lambda : \mathbb{X} \to \mathbb{X}$ is an operator defined by* (3.11) *and having property (**A**). Let u_* be the solution of* (2.1) *and u_Λ the solution of* (2.8)*. If $\lambda_+/c_\Lambda = \mathcal{O}(1)$ as $\lambda_+ \to 0$, then u_Λ converges strongly to u_*, that is,*

$$\|u_\Lambda - u_*\| \to 0, \quad \text{as} \quad \lambda_+ \to 0.$$

Proof. It follows from Theorem 3.1 that the error of the regularization solution $e_\Lambda := u_* - u_\Lambda$ converges weakly to zero as $\lambda_+ \to 0$. We prove that in this case the convergence is strong. In fact, it follows from (2.1) and (2.8) that

$$(\Lambda + \mathcal{K}^*\mathcal{K})e_\Lambda = \Lambda u_*.$$

By property (**A**), the above equation and (3.11), we obtain the inequality

$$c_\Lambda\|e_\Lambda\|^2 \leq ((\Lambda + \mathcal{K}^*\mathcal{K})e_\Lambda, e_\Lambda) = (\Lambda u_*, e_\Lambda) = \sum_{j\in\mathbb{Z}_{N+1}} \lambda_j(\mathcal{Q}_j u_*, e_\Lambda),$$

which leads to the estimate

$$\|e_\Lambda\|^2 \leq \frac{1}{c_\Lambda} \sum_{j\in\mathbb{Z}_{N+1}} \lambda_j(\mathcal{Q}_j u_*, e_\Lambda).$$

Since $0 \leq \lambda_j \leq \lambda_+$, for all $j \in \mathbb{Z}_{N+1}$, we conclude that

$$\|e_\Lambda\|^2 \leq \frac{\lambda_+}{c_\Lambda} \sum_{j\in\mathbb{Z}_{N+1}} |(\mathcal{Q}_j u_*, e_\Lambda)|. \tag{3.12}$$

Noting that the integer N is fixed, in (3.12) letting $\lambda_+ \to 0$, using the hypothesis on the ratio of λ_+ and c_Λ, and employing the weak convergence proved in Theorem 3.1, we see that $\|e_\Lambda\| \to 0$ as $\lambda_+ \to 0$.

We now turn to establishing a convergence result for the regularization methods (2.8) and (2.9) when $\mathbb{X} = \mathbb{Y}$ and Λ has an additional property which we define next. Operator Λ is said to have property (**B**) if it is a self-adjoint bounded and positive semi-definite linear operator and it commutes with $\mathcal{K}$, i.e., $\Lambda\mathcal{K} = \mathcal{K}\Lambda$. An example of such operators is $\Lambda := f(\mathcal{K})$, where $f(x)$ is a polynomial in x. However, the regularization operator Λ in this case is not necessarily in the form (2.7).

Lemma 3.1. *Suppose that $\mathcal{K}$ is a linear injective compact operator from $\mathbb{X}$ to $\mathbb{X}$, and $\Lambda : \mathbb{X} \to \mathbb{X}$ is an operator having properties (**A**) and (**B**). Then*

$$\|(\Lambda + \mathcal{K}^*\mathcal{K})^{-1}\Lambda\| \leq 1, \tag{3.13}$$

$$\|(\Lambda + \mathcal{K}^*\mathcal{K})^{-1}\mathcal{K}^*\mathcal{K}\| \leq 1 \tag{3.14}$$

and

$$\|(\Lambda + \mathcal{K}^*\mathcal{K})^{-1}\mathcal{K}^*\| \leq 1/\sqrt{c_\Lambda}. \tag{3.15}$$

Proof. By using the property (**B**), we have that

$$\begin{aligned} \|(\Lambda + \mathcal{K}^*\mathcal{K})x\|^2 &= \|\Lambda x\|^2 + \|\mathcal{K}^*\mathcal{K}x\|^2 + 2(\Lambda x, \mathcal{K}^*\mathcal{K}x) \\ &= \|\Lambda x\|^2 + \|\mathcal{K}^*\mathcal{K}x\|^2 + 2(\Lambda\mathcal{K}x, \mathcal{K}x). \end{aligned}$$

The positive semi-definiteness of Λ ensures that

$$\|(\Lambda + \mathcal{K}^*\mathcal{K})x\| \geq \|\Lambda x\| \quad \text{and} \quad \|(\Lambda + \mathcal{K}^*\mathcal{K})x\| \geq \|\mathcal{K}^*\mathcal{K}x\|.$$

It follows from property (**A**) that $(\Lambda + \mathcal{K}^*\mathcal{K})^{-1}$ exists. Thus, the above two inequalities imply that

$$\|\Lambda(\Lambda + \mathcal{K}^*\mathcal{K})^{-1}\| \leq 1 \quad \text{and} \quad \|\mathcal{K}^*\mathcal{K}(\Lambda + \mathcal{K}^*\mathcal{K})^{-1}\| \leq 1. \tag{3.16}$$

From property (**B**) we have that

$$(\Lambda + \mathcal{K}^*\mathcal{K})^{-1}\Lambda = \Lambda(\Lambda + \mathcal{K}\mathcal{K}^*)^{-1} \quad \text{and} \quad (\Lambda + \mathcal{K}^*\mathcal{K})^{-1}\mathcal{K}^*\mathcal{K} = \mathcal{K}^*\mathcal{K}(\Lambda + \mathcal{K}\mathcal{K}^*)^{-1}.$$

Inequalities (3.13) and (3.14) follows directly from estimates (3.16) and the above two equations.

It remains to prove the estimate (3.15). A way similar to the proof of the second estimate in (3.16) leads to

$$\|\mathcal{K}\mathcal{K}^*(\Lambda + \mathcal{K}\mathcal{K}^*)^{-1}\| \leq 1. \tag{3.17}$$

From property (**B**) we also have that

$$(\Lambda + \mathcal{K}^*\mathcal{K})^{-1}\mathcal{K}^* = \mathcal{K}^*(\Lambda + \mathcal{K}\mathcal{K}^*)^{-1}.$$

This ensures that for any $x \in \mathbb{X}$,

$$\begin{aligned} \|(\Lambda + \mathcal{K}^*\mathcal{K})^{-1}\mathcal{K}^*x\|^2 &= (\mathcal{K}^*(\Lambda + \mathcal{K}\mathcal{K}^*)^{-1}x, \mathcal{K}^*(\Lambda + \mathcal{K}\mathcal{K}^*)^{-1}x) \\ &= (\mathcal{K}\mathcal{K}^*(\Lambda + \mathcal{K}\mathcal{K}^*)^{-1}x, (\Lambda + \mathcal{K}\mathcal{K}^*)^{-1}x) \\ &\leq \|\mathcal{K}\mathcal{K}^*(\Lambda + \mathcal{K}\mathcal{K}^*)^{-1}\|\|(\Lambda + \mathcal{K}\mathcal{K}^*)^{-1}\|\|x\|^2. \end{aligned}$$

Using (3.17) in the above inequality yields

$$\|(\Lambda + \mathcal{K}^*\mathcal{K})^{-1}\mathcal{K}^*x\|^2 \le \|(\Lambda + \mathcal{K}\mathcal{K}^*)^{-1}\|\|x\|^2.$$

Combining this inequality with (2.11) gives the estimate (3.15).

The above lemma leads to the next convergence theorem.

Theorem 3.3. *Suppose that $\mathcal{K}$ is a linear injective compact operator from $\mathbb{X}$ to $\mathbb{X}$, and that Λ is an operator having properties (**A**) and (**B**). Let u_*, u_Λ and u_Λ^δ be the solutions of* (2.1), (2.8) *and* (2.9), *respectively. Then*

$$\|u_* - u_\Lambda\| \to 0, \ \textit{as} \ \|\Lambda\| \to 0,$$

and

$$\|u_* - u_\Lambda^\delta\| \to 0, \quad \text{as} \quad \|\Lambda\| \to 0, \quad \delta \to 0 \quad \text{and} \quad \delta/\sqrt{c_\Lambda} \to 0.$$

Proof. It follows from the second inequality of Lemma 3.1 that

$$\|u_\Lambda\| = \|(\Lambda + \mathcal{K}^*\mathcal{K})^{-1}\mathcal{K}^*\mathcal{K}u_*\| \le \|u_*\|. \tag{3.18}$$

Thus, there exist a subsequence u_{Λ_j} of u_Λ and an element $v \in \mathbb{X}$ such that

$$u_{\Lambda_j} \rightharpoonup v, \quad \text{with} \quad \|\Lambda_j\| \to 0, \quad \text{as} \quad j \to \infty. \tag{3.19}$$

From the equation (2.8), (3.19) and the injectivity of the operator $\mathcal{K}$ we conclude $v = u_*$. This with (3.19) and (3.18) yields

$$\|u_*\| \le \liminf_{j\to\infty} \|u_{\Lambda_j}\| \le \limsup_{j\to\infty} \|u_{\Lambda_j}\| \le \|u_*\|,$$

which leads to $\lim_{\|\Lambda\|\to 0} \|u_\Lambda\| = \|u_*\|$, from which with (3.19) we obtain the first desired convergence result.

To prove the second convergence result, we use

$$\|u_* - u_\Lambda^\delta\| \le \|u_* - u_\Lambda\| + \|u_\Lambda - u_\Lambda^\delta\|. \tag{3.20}$$

By applying the third inequality of Lemma 3.1, we find that

$$\|u_\Lambda - u_\Lambda^\delta\| = \|(\Lambda + \mathcal{K}^*\mathcal{K})^{-1}\mathcal{K}^*(f - f^\delta)\| \le \delta/\sqrt{c_\Lambda}. \tag{3.21}$$

The second convergence follows from inequality (3.20), this estimate and the first convergence of this theorem.

The convergence established in the last theorem holds as $\|\Lambda\| \to 0$. Although $\|\Lambda\|$ might be viewed as one "parameter", the operator Λ can contain multiple parameters which encode different noise scales. Hence, it significantly differs from the single parameter regularization. Along this line, we present in the following theorem an error estimate for the regularization solution u_Λ^δ defined by (2.9) when Λ has property (**B**). To prepare for the proof of this error estimate, we first recall in the next lemma an interpolation inequality for the positive semi-definite operator $\mathcal{A}$, which can be found in [19]. For convenience of the reader, we provide a proof of this inequality.

Lemma 3.2. *If $0 \le \nu \le 1$, then there exists a positive constant c such that for all positive semi-definite operators $\mathcal{A}$ and all $x \in \mathbb{X}$,*

$$\|\mathcal{A}^\nu x\| \le c\|\mathcal{A}x\|^\nu \|x\|^{1-\nu}.$$

Proof. It is clear that the inequality holds if $\nu = 0, 1$ or $x = 0$. It remains to prove the inequality when $0 < \nu < 1$ and $x \ne 0$. Note that

$$\mathcal{A}^\nu x = \frac{\sin \pi\nu}{\pi} \int_0^{+\infty} s^{\nu-1}(s\mathcal{I} + \mathcal{A})^{-1}\mathcal{A}x \mathrm{d}s.$$

We choose $\eta := \|\mathcal{A}x\|/\|x\|$ and write

$$\mathcal{A}^\nu x = \frac{\sin \pi\nu}{\pi} \int_0^{\eta} s^{\nu-1}(s\mathcal{I} + \mathcal{A})^{-1}\mathcal{A}x \mathrm{d}s + \frac{\sin \pi\nu}{\pi} \int_\eta^{+\infty} s^{\nu-1}(s\mathcal{I} + \mathcal{A})^{-1}\mathcal{A}x \mathrm{d}s.$$

By using $\|(s\mathcal{I} + \mathcal{A})^{-1}\mathcal{A}\| \le 1$ and $\|(s\mathcal{I} + \mathcal{A})^{-1}\| \le 1/s$, we obtain that

$$\begin{aligned}\|\mathcal{A}^\nu x\| &\le \frac{|\sin \pi\nu|}{\pi} \int_0^{\eta} s^{\nu-1}\mathrm{d}s\|x\| + \frac{|\sin \pi\nu|}{\pi} \int_\eta^{+\infty} s^{\nu-2}\mathrm{d}s\|\mathcal{A}x\| \\ &\le \Big(\frac{|\sin \pi\nu|}{\nu\pi} + \frac{|\sin \pi\nu|}{(1-\nu)\pi}\Big)\|\mathcal{A}x\|^\nu\|x\|^{1-\nu},\end{aligned}$$

which proves the result of this lemma.

We are now ready to prove the theorem.

Theorem 3.4. *Suppose that $\mathcal{K}$ is a linear injective compact operator from $\mathbb{X}$ to $\mathbb{X}$, and $\Lambda : \mathbb{X} \to \mathbb{X}$ has properties (**A**) and (**B**). Let u_* and u_Λ^δ be the solutions of* (2.1) *and* (2.9)*, respectively. If $u_* = (\mathcal{K}^*\mathcal{K})^\nu\omega$ with $\omega \in \mathbb{X}$, and $0 < \nu \le 1$, then there exists a constant c independent of Λ and δ such that*

$$\|u_* - u_\Lambda^\delta\| \le c\|\omega\|\|\Lambda\|^\nu + \delta/\sqrt{c_\Lambda}.$$

Proof. Since $u_* = (\mathcal{K}^*\mathcal{K})^\nu\omega$, we have that

$$u_* - u_\Lambda = (\Lambda + \mathcal{K}^*\mathcal{K})^{-1}\Lambda(\mathcal{K}^*\mathcal{K})^\nu\omega.$$

This with property (**B**) yields that

$$u_* - u_\Lambda = [(\Lambda + \mathcal{K}^*\mathcal{K})^{-1}\Lambda]^{(1-\nu)}[(\Lambda + \mathcal{K}^*\mathcal{K})^{-1}\mathcal{K}^*\mathcal{K}]^\nu\Lambda^\nu\omega.$$

By repeatedly using Lemma 3.2, we find that there exists a positive constant c such that for all positive semi-definite operators Λ

$$\|u_* - u_\Lambda\| \le c\|(\Lambda + \mathcal{K}^*\mathcal{K})^{-1}\Lambda\|^{(1-\nu)}\|(\Lambda + \mathcal{K}^*\mathcal{K})^{-1}\mathcal{K}^*\mathcal{K}\|^\nu\|\Lambda\|^\nu\|\omega\|.$$

By employing the first and second estimates of Lemma 3.1, we obtain that $\|u_* - u_\Lambda\| \le c\|\omega\|\|\Lambda\|^\nu$. Combining (3.20), (3.21) and above inequality yields the desired estimate.

In the remainder of this section we present a special result when the space $\mathbb{W}_i$, $i \in \mathbb{N}_0$, that appear in (2.4) are chosen to be the eigenfunction spaces of the operator $\mathcal{K}^*\mathcal{K}$. Choosing the spaces $\mathbb{W}_i$, $i \in \mathbb{N}_0$, as the eigenfunction spaces of the operator $\mathcal{K}^*\mathcal{K}$ is practical for computation purpose only if the eigenvalues and eigenfunctions are available, which are possible in some cases. Nevertheless, they give intrinsic properties of the multi-parameter regularization for ill-posed equations.

Suppose that $\mathcal{K} : \mathbb{X} \to \mathbb{Y}$ is injective and $\{\phi_i, \psi_i, \mu_i : i \in \mathbb{N}\}$ is the singular system of $\mathcal{K}$. Thus, $\{\phi_n : n \in \mathbb{N}\}$ is a complete orthonormal system of the space $\mathbb{X}$. We remark that the classical Picard theorem for single regularization solution can be extended to the multi-parameter regularization solution in the form

$$u_\Lambda = \sum_{n\in\mathbb{N}} \mu_n(f, \psi_n)(\Lambda + \mu_n^2\mathcal{I})^{-1}\phi_n \tag{3.22}$$

and

$$u_\Lambda^\delta = \sum_{n\in\mathbb{N}} \mu_n(f^\delta, \psi_n)(\Lambda + \mu_n^2\mathcal{I})^{-1}\phi_n. \tag{3.23}$$

For $n \in \mathbb{N}_0$, let $\mathbb{W}_n^0 := \text{span}\{\phi_n\}$, $\mathcal{Q}_n^0$ be the orthogonal projection from $\mathbb{X}$ onto $\mathbb{W}_n^0$ and Λ_0 be the regularization operator corresponding to $\mathcal{Q}_n^0$. In this setting, equation (2.8) becomes

$$(\Lambda_0 + \mathcal{K}^*\mathcal{K})u_{\Lambda_0} = \mathcal{K}^* f \tag{3.24}$$

and equation (2.9) becomes

$$(\Lambda_0 + \mathcal{K}^*\mathcal{K})u_{\Lambda_0}^\delta = \mathcal{K}^* f^\delta. \tag{3.25}$$

To present the error of the regularization solution, in addition to the hypothesis $\|f - f^\delta\| \le \delta$ for some $\delta > 0$ as we always assume in the single parameter regularization, we suppose that the noise in the given data f has the representation that

$$f - f^\delta = \sum_{n\in\mathbb{N}_0} \eta_n\psi_n, \quad \text{with} \quad |\eta_n| \le \delta_n, \quad \delta = \left(\sum_{n\in\mathbb{N}_0} \delta_n^2\right)^{\frac{1}{2}}, \tag{3.26}$$

where $\psi_0 \in N(\mathcal{K}^*)$, the null space of $\mathcal{K}^*$, with $\|\psi_0\| = 1$. As direct consequences of (3.22) and (3.23), we have the next result.

Proposition 3.1. *Suppose that $\mathcal{K}$ is a linear injective compact operator from $\mathbb{X}$ to $\mathbb{Y}$ and that the operator Λ_0 is chosen so that the inverse operator $(\Lambda_0 + \mathcal{K}^*\mathcal{K})^{-1}$ exists. Let u_{Λ_0} and $u_{\Lambda_0}^\delta$ be the solution of equation* (3.25) *and* (3.24), *respectively. Then*

$$\|u_* - u_{\Lambda_0}\|^2 = \sum_{n\in\mathbb{N}} |(u_*, \phi_n)|^2 \left(\frac{\lambda_n}{\lambda_n + \mu_n^2}\right)^2, \tag{3.27}$$

and

$$\|u_* - u_{\Lambda_0}^\delta\|^2 = \sum_{n\in\mathbb{N}} \left[\frac{\lambda_n(u_*, \phi_n) + \eta_n\mu_n}{\lambda_n + \mu_n^2}\right]^2. \tag{3.28}$$

4. Parameter Choices and Numerical Experiments

A crucial and challenging issue for the regularization method is the choice of regularization parameters. The general principle for the choice of parameters is to minimize the approximation error while controlling the condition number of the resulting regularized operator or matrix. The choices of parameters for the multi-parameter regularization method require further substantial study to provide sound theoretical results. However, in this section we discuss several practical parameter choice strategies for the matrix case, based on different a prior knowledge of the original matrix. The ideas for these strategies are motivated by the single parameter regularization (cf. [11, 15, 20, 21, 24, 25]). We point out that once multi-parameters are chosen as suggested by a choice strategy, the computational cost for the multi-parameter regularization is the same as that for the single parameter regularization. We also provide examples to compare the numerical performance of the multi-parameter regularization via the single parameter regularization. For convenience, we use only finite dimensional examples.

The first parameter choice strategy is for the special case when the eigenfunction spaces are used for the decomposition of the space $\mathbb{X}$ and f^δ satisfies (3.26). Equation (3.28) in Proposition 3.1 leads to the parameter choice strategy by minimizing the right hand side of (3.28) to obtain parameters λ_n. Specifically, for $n \in \mathbb{N}$, we define the function $\Phi_n(\lambda) := \left[\frac{\lambda(u_*,\phi_n)+\eta_n\mu_n}{\lambda+\mu_n^2}\right]^2$ and choose $\lambda_n := \arg\min_{\lambda\geq 0} \Phi_n(\lambda)$. This results in the following rule for choices of parameters.

Rule 1: For each $n \in \mathbb{N}$ choose λ_n as follows.

1. If $(u_*, \phi_n)\eta_n < 0$, then $\lambda_n := -\eta_n\mu_n/(u_*, \phi_n)$.
2. Suppose $(u_*, \phi_n)\eta_n > 0$. If $|(u_*, \phi_n)|\mu_n > |\eta_n|$, then $\lambda_n := 0$; if $|(u_*, \phi_n)|\mu_n < |\eta_n|$, then $\lambda_n := +\infty$; if $(u_*, \phi_n)\mu_n = \eta_n$, then $\lambda_n \geq 0$.
3. If $(u_*, \phi_n) = 0$, then $\lambda_n := +\infty$.

Note that Rule 1 is an ideal a priori parameter choice strategy which uses the information on u_*. It can not be implemented in practice. For this reason, we modify Rule 1 to propose an a posteriori parameter choice strategy. In fact, noting that $|(u_*, \phi_n)|\mu_n = |(f, \psi_n)|$, we replace $|(u_*, \phi_n)|\mu_n$ by $|(f^\delta, \psi_n)|$, which approximates $|(f, \psi_n)|$ and is also computable. Moreover, we replace η_n by its bound δ_n which we assume to be available. This leads to the next implementable rule for choices of parameters.

Rule 2: For each $n \in \mathbb{N}$, choose λ_n as follows.

1. If $|(f^\delta, \psi_n)| \leq \delta_n$, then $\lambda_n := +\infty$.
2. If $|(f^\delta, \psi_n)| > \delta_n$, then $\lambda_n := \mu_n^2\delta_n/(|(f^\delta, \psi_n)| - \delta_n)$.

We present the first numerical example.

Example 1. Consider solving the linear system

$$Ku = f, \quad \text{where} \quad K = \begin{bmatrix} 10 & 7 & 8 & 7 \\ 7 & 5 & 6 & 5 \\ 8 & 6 & 10 & 9 \\ 7 & 5 & 9 & 10 \end{bmatrix}, \quad f = \begin{bmatrix} 32 \\ 23 \\ 33 \\ 31 \end{bmatrix}. \tag{4.1}$$

The exact solution of equation (4.1) is given by $u_* = [1, 1, 1, 1]^T$. The matrix K has the singular value decomposition $K = U\,\text{diag}(\mu_i)\,V^T$, where $U = (u_1, u_2, u_3, u_4)$, $V = (v_1, v_2, v_3, v_4)$, $\mu_1 = 30.29, \mu_2 = 3.86, \mu_3 = 0.84, \mu_4 = 0.010$. We then find that the condition number of matrix K is $\text{cond}(K) := \frac{\mu_1}{\mu_4} = 2984.09$. Hence the linear system is ill-conditioned. In fact, if the right hand side f is perturbed by noise to the vector

$$f^\delta := (32.1343, 23.0039, 33.1249, 30.9204)^T,$$

then the solution of the corresponding equation is

$$u^\delta = \mathcal{K}^{-1} f^\delta = (5.9234, -7.1591, 3.1397, -0.3005)^T.$$

Since the error of this approximate solution is $\|u^\delta - u_*\| = 9.8529$, it is not an acceptable approximate solution. Thus, we have to apply a regularization method to solve it.

To use the above multi-parameter regularization method we assume that $f - f^\delta = \sum_{j=1}^4 \eta_j u_j$ and $\|f - f^\delta\| \le \delta$, with $|\eta_n| \le \delta_n, n = 1, 2, 3, 4$, where $\|\cdot\|$ is the Euclidean norm, and $\delta^2 = \sum_{j=1}^4 \delta_j^2$. The regularization solution has the form

$$u_{\Lambda_0}^\delta = \sum_{i=1}^4 u_i^T f^\delta \frac{\mu_i}{\lambda_i + \mu_i^2} v_i. \tag{4.2}$$

In the numerical results presented below we choose noise $\delta = 0.2$ with different decompositions $\eta := (\eta_1, \eta_2, \eta_3, \eta_4)$ and $\delta := (\delta_1, \delta_2, \delta_3, \delta_4)$ which satisfy $\delta_n = |\eta_n|$ for $n = 1, 2, 3, 4$:

Case 1 :	$\eta = (0.1, 0.1, 0.1, 0.1)$;	Case 2 :	$\eta = (0.1, 0.1, 0.1, -0.1)$;
Case 3 :	$\eta = (0.1, 0.1, -0.1, -0.1)$;	Case 4 :	$\eta = (-0.1, -0.1, 0.1, 0.1)$;
Case 5 :	$\eta = (-0.1, 0.1, -0.1, 0.1)$;	Case 6 :	$\eta = (\sqrt{0.2}, \sqrt{0.2}, 0, 0)$;
Case 7 :	$\eta = (-\sqrt{0.2}, \sqrt{0.2}, 0, 0)$;	Case 8 :	$\eta = (0, \sqrt{0.2}, \sqrt{0.2}, 0)$;
Case 9 :	$\eta = (0, 0, \sqrt{0.2}, \sqrt{0.2})$;	Case 10 :	$\eta = (0, 0, \sqrt{0.2}, -\sqrt{0.2})$.

For comparison, we present a numerical result for the single parameter regularization with optimal parameter $\alpha_0 := \arg\min_{\alpha>0} \|u_\alpha^\delta - u_*\|$. In Table 4.1, we list the parameters Λ_0 and Λ_0' obtained by Rules 1 and 2, respectively, and in Table 4.2, we present the computed errors of the regularization solutions corresponding to these choices. These numerical results show that the multi-parameter regularization provides more accurate regularization solutions than the single parameter regularization and the a posteriori parameter choice strategy (Rule 2) works as well as Rule 1. We also observe that the regularization parameters are sensitive to the noise decomposition and thus, they affect the result of regularization.

Next we present three practical choices of multi-parameters.

Strategy one is based on the assumption that spectrums of matrix $\mathcal{K}^*\mathcal{K}$ have a decay trend along its diagonal blocks in a multi-scale representation. We aim at reducing the condition number of the resulting regularized matrix by choosing appropriate parameters. To this end, we balance the distribution of the spectrums along the scales. Specifically, for each $i \in \mathbb{Z}_{N+1}$ we compute the spectrums of $\mathcal{Q}_i\mathcal{K}^*\mathcal{K}\mathcal{Q}_i$ and their mean value σ_i. We then set $\sigma_{max} := \max\{\sigma_i : i \in \mathbb{Z}_{N+1}\}$, and choose $\lambda_i := \sigma_{max} - \sigma_i$, for $i \in \mathbb{Z}_{N+1}$. Since by hypothesis, the absolute values of entries decay quickly along the diagonal blocks, by the Weyl theorem, this choice of parameters λ_i will result in a regularized matrix having condition number close to 1. This choice was used in [13] in a numerical experiment for signal processing which gives satisfactory results.

Strategy two uses a priori knowledge on the solution and noise. Recall that in the single parameter case, following [15] a natural choice of regularization parameter is the noise-solution ratio, that is, $\alpha := \|f - f^\delta\|^2/\|u\|^2$. We now extend this idea to the multiple parameter setting. For given numbers δ and $\mathbf{q} := (q_1, q_2, \cdots, q_N)$, we call $u_* \in \mathbb{X}$ a generalized quasi-solution of the equation (2.1) if it satisfies

$$\|\mathcal{K}u_* - f\| \le \delta, \quad \|\mathcal{Q}_i u_*\| \le q_i, \qquad i \in \mathbb{Z}_{N+1}. \tag{4.3}$$

Table 4.1: Parameter choices.

Case	α_0	$\Lambda_0 = (\lambda_1, \lambda_2, \lambda_3, \lambda_4)$	$\Lambda_0' = (\lambda_1', \lambda_2', \lambda_3', \lambda_4')$
1	5.150	(1.528, 3.379, $+\infty$, $+\infty$)	(1.528, 3.379, $+\infty$, $+\infty$)
2	4.155E-3	(1.528, 3.379, $+\infty$, 4.159E-3)	(1.528, 3.379, $+\infty$, 4.159E-3)
3	4.156E-3	(1.528, 3.379, 5.381, 4.159E-3)	(1.528, 3.379, 5.381, 4.159E-3)
4	3.196	(0, 0, $+\infty$, $+\infty$)	(1.534, 6.190, $+\infty$, $+\infty$)
5	3.833	(0, 3.379, 5.381, $+\infty$)	(1.534, 3.379, 5.381, $+\infty$)
6	0	(6.835, 1.511E+1, 0, 0)	(6.835, 1.511E+1, 0, 0)
7	0	(0, 1.511E+1, 0, 0)	(6.939, 1.511E+1, 0, 0)
8	1.475E+1	(0, 1.511E+1, $+\infty$, 0)	(0, 1.511E+1, $+\infty$, 0)
9	9.975	(0, 0, $+\infty$, $+\infty$)	(0, 0, $+\infty$,$+\infty$)
10	9.465	(0, 0, $+\infty$, 1.86E-2)	(0, 0, $+\infty$, 1.86E-2)

Table 4.2: Computed error estimates.

Case	$\|u^\delta_{\alpha_0} - u\|$	$\|u^\delta_{\Lambda_0} - u\|$	$\|u^\delta_{\Lambda_0'} - u\|$	Case	$\|u^\delta_{\alpha_0} - u\|$	$\|u^\delta_{\Lambda_0} - u\|$	$\|u^\delta_{\Lambda_0'} - u\|$
1	2.462E-1	2.445E-1	2.445E-1	6	1.169E-1	6.205E-14	6.205E-14
2	1.209E-1	1.567E-2	1.567E-2	7	1.169E-1	1.477E-2	2.953E-2
3	1.207E-1	4.682E-14	4.682E-14	8	2.492E-1	1.567E-2	1.567E-2
4	2.504E-1	2.459E-1	2.500E-1	9	2.546E-1	2.445E-1	2.445E-1
5	2.446E-1	2.440E-1	2.441E-1	10	2.537E-1	1.567E-2	1.567E-2

We assume that problem (4.3) is stable in the sense that

$$S(\delta, \mathbf{q}) := \sup\{\|u\| : u \in \mathbb{X}, \|\mathcal{K}u\| \le \delta \text{ and } \|\mathcal{Q}_i u\| \le q_i \text{ for } i \in \mathbb{Z}_{N+1}\} \to 0, \quad \text{as} \quad \delta \to 0.$$

We also assume that the noise distribution on different scales is known, that is, there is a vector $(\delta_0, \delta_1, \cdots, \delta_N)$ such that

$$\sum_{i \in \mathbb{Z}_{N+1}} \delta_i^2 = \delta^2, \quad \|\mathcal{Q}_i(f - f^\delta)\| \le \delta_i, \quad i \in \mathbb{Z}_{N+1}.$$

Note that δ_i is an upper bound to control the noise in scale space $\mathbb{W}_i$. We choose $\lambda := (\lambda_0, \lambda_1, \cdots, \lambda_N)$ with

$$\lambda_i = \delta_i^2 / q_i^2, \qquad i \in \mathbb{Z}_{N+1}. \tag{4.4}$$

From (2.12), solving the multi-parameter regularization equation is equivalent to minimizing the functional $F(u)$. It can be shown that if λ_i, $i \in \mathbb{Z}_{N+1}$, are chosen as in (4.4) and if $u \in \mathbb{X}$ satisfies the condition

$$\|\mathcal{K}u - f\|^2 + \sum_{i \in \mathbb{Z}_{N+1}} \lambda_i \|\mathcal{Q}_i u\|^2 \le 2\delta^2,$$

then it is a nice approximation of the generalized quasi-solution u_* that satisfies (4.3).

In the next result we present the condition number of the regularized operator defined by $\mathcal{K}_\Lambda := \mathcal{K}^*\mathcal{K} + \sum_{i \in \mathbb{Z}_{N+1}} \lambda_i \mathcal{Q}_i$, using the choice of parameters. We define $\underline{\delta}$ and $\overline{\delta}$ to be the minimum and maximum values of the components of δ, respectively, and define $\underline{q}$ and $\overline{q}$ likewise. If λ_i, $i \in \mathbb{Z}_{N+1}$, are chosen as in (4.4) and if $\mathcal{Q}_i$ are orthogonal projections, then

$$\text{cond}\,(\mathcal{K}_\Lambda) \le [(\overline{q}\underline{q}\|\mathcal{K}\|)^2 + (\overline{q}\overline{\delta})^2]/\underline{\delta}^2.$$

Strategy three is proposed for regularization of an $n \times n$ matrix K which has a special structure in which an orthonormal basis of the null space $N(K)$ of K is easy to obtain. In particular, when K is of the Γ-type in the sense that the matrix can be partitioned into four blocks, where the lower right block is the zero square matrix and its order is much larger than that of the upper left block, the matrix has the above structure. The (compressed) matrix representation for an integral operator with a smooth kernel in a multi-scale basis is of the Γ-type. Let B^t be the matrix whose column vectors form an orthonormal basis of the null space $N(K)$. Set $m = \dim N(K)$ and let $\mathbb{S}$ denote the orthogonal complement of $N(K)$ in $\mathbb{R}^n$. Then the dimension of $\mathbb{S}$ is $n-m$. We denote by J the matrix representation of the orthogonal projection from $\mathbb{R}^n$ onto $\mathbb{S}$. It can be shown that $N(K) = N(K^tK)$. Let σ_i, $i = 1, 2, \cdots, n-m$, denote the nonzero eigenvalues of K^tK. Let $\overline{\sigma}$ and $\underline{\sigma}$ be the maximal and minimal eigenvalue of K^tK, respectively. We choose

$$\lambda_0 = \overline{\sigma}/2 - \underline{\sigma}, \quad \lambda_i = \overline{\sigma}/2, \quad i = 1, 2, \cdots, N. \tag{4.5}$$

We define a block diagonal matrix $\Lambda := \mathrm{diag}(\lambda_0 I_0, \lambda_1 I_1, \cdots, \lambda_N I_N)$, where I_i are identity matrices and the order of I_1 is equal to $n-m$, the dimension of $\mathbb{S}$. We propose a regularization method by

$$\left(K^tK + (J^t, B^t)\Lambda(J^t, B^t)^t\right) u_\Lambda = K^t f.$$

Note that in this method we choose the regularization matrix as $(J^t, B^t)\Lambda(J^t, B^t)^t$ which is not necessarily a diagonal matrix. It can be proved that if λ_i, $i \in \mathbb{Z}_{N+1}$, are chosen as in (4.5), the condition number of the matrix $K^tK + (J^t, B^t)\Lambda(J^t, B^t)^t$ is bounded above by 3.

In the following, we present three numerical experiments in signal processing. The first experiment deals with noise uniformly distributed in the frequency domain, the second considers noise not uniformly distributed in the frequency domain and the third treats noise piecewise distributed in the time domain. All of these examples demonstrate that the multi-parameter regularization performs better than the single parameter regularization.

Example 2. Consider solving the linear system

$$Fu = h, \tag{4.6}$$

where F is a singular symmetric toeplitz matrix. Setting $\tilde{K} := F^T F$ and $f := F^t h$, using the least squares method, the linear system becomes

$$\tilde{K} u = f. \tag{4.7}$$

We suppose that the data f contains the Gaussian white noise which is assumed to be uniformly distributed in the frequency domain. In order to obtain a multiscale representation for $\tilde{K}$, we use the low-pass filter $\{a, b, c, d\}$ and the high-pass filter $\{d, -c, b, -a\}$, with

$$a := \frac{1+\sqrt{3}}{4\sqrt{2}}, \quad b := \frac{3+\sqrt{3}}{4\sqrt{2}}, \quad c := \frac{3-\sqrt{3}}{4\sqrt{2}}, \quad d := \frac{1-\sqrt{3}}{4\sqrt{2}},$$

to form the wavelet transform matrix $P_{l,n}$, where l indicates the number of high frequency levels used and 2^n denotes the length of data, see [13]. The transform converts $\tilde{K}$ into $K := P_{l,n}\tilde{K}P_{l,n}$ having a multiscale structure. For positive λ_i, $i = 1, 2, \cdots, \ell+1$, we introduce the regularization matrix

$$\Lambda := \mathrm{diag}\ (\lambda_1 I_{2^{n_0}}, \lambda_2 I_{2^{n_0}}, \lambda_3 I_{2^{n_0+1}}, \cdots, \lambda_{\ell+1} I_{2^{n-1}}), \tag{4.8}$$

where 2^{n_0} denotes the size of the block corresponding to the coarsest level. The multi-parameter regularization equation is then given by

$$(\Lambda + K)P_{\ell,n} u_\Lambda = P_{\ell,n} g. \tag{4.9}$$

Table 4.3: The square norm of noise and data on each band.

	Band 4	Band 3	Band 2	Band 1
The square norm of noise	7.8479	7.3132	11.5803	16.0512
The relative square norm of noise	0.3531	0.3290	0.5210	0.7221
The square norm of signal	405.5445	67.6911	32.9987	30.4510
The relative square norm of signal	0.9967	0.1664	0.0811	0.0748
SNR	0.0194	0.1080	0.3509	0.5271

In our simulation, we choose F to be the symmetric toeplitz singular matrix with the first row given by $[0.1467, 0.0962, 0.0267, 0.003, 0.0001, 0, \cdots, 0]$. The original "Piece-Regular" data u in our experiment is taken from the WaveLab toolbox at "www.stat.stanford.edu/˜wavelab/". We choose the length of data as 512 and add to data h the Gaussian noise with zero mean and variance $\sigma = 1$. We perform three times wavelet transforms on the noise and data, which correspond to $\ell = 3$ and decompose all functions to four frequency bands. In Table 4.3, we report the square norm of noise and data on different bands and the corresponding signal-to-noise ratio (SNR). From Table 4.3, we find that the Gaussian white noise has a scaling relation among different bands, that is, $\delta_i \approx 2^{-\frac{i}{2}}\delta$, where δ_i is the square norm of noise at the i-th band and δ is the square norm of noise. However, we do not find the same relation for a practical data, which has a multi-scale structure but is not simply scaling.

We choose the regularization parameters $\lambda_1 = 0.0194$, $\lambda_2 = 0.1080$, $\lambda_3 = 0.3509$, $\lambda_4 = 0.5271$ according to strategy two. Clearly, they do not have the same scaling relation as the noise components do, due to the non-scaling relation of the signal. We build the multi-parameter regularization matrix Λ by (4.8) accordingly with $\ell := 3$ and $n_0 := 6$, and solve numerically the multi-parameter regularization equation (4.9) with $n := 9$. In Fig. 4.1 we compare the restored signals by using the multi-parameter regularization and the single parameter regularization. In the single parameter regularization model we use the optimal parameter $\alpha = 0.08$. Clearly, the result obtained from using the multi-parameter regularization is significantly better than that from using the single parameter regularization.

Example 3. Consider a denoising problem in petroleum industry. In petroleum drilling, an electric detonator is used to detect the drilling curve. As a pre-process, one has to recover the original data from noisy data, where the noise is mainly from a nearby electricity device. We formulate the denoising problem as an ill-posed problem described by

$$f = Ku + w, \tag{4.10}$$

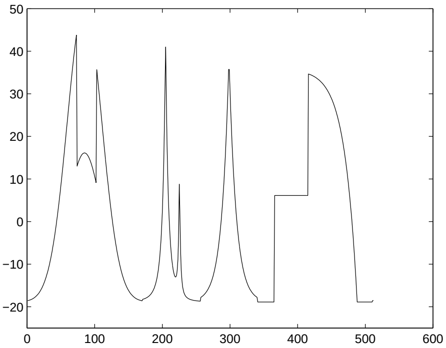
where f is the observed data, w is noise and K is the identity matrix. Since we already know that the electric device adds a strong 50HZ noise to the recorded data, we propose a regularization method according to this special situation.

In our numerical test, we fix the length of data as 1024. Because of the previous knowledge of noise, we consider a decomposition of $\mathbb{R}^{1000}$ as

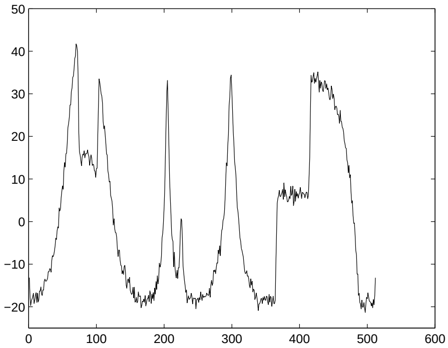
$$\mathbb{R}^{1000} = \bigoplus_{i=1}^{1000} \mathbb{W}_i,$$

where

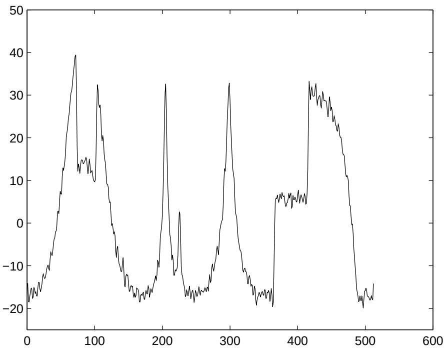
$$\mathbb{W}_i = \text{ span } \{e_i\}, \text{ with } e_i(j) = \cos\frac{\pi(2j-1)(i-1)}{2\times 1000}, \quad j = 1, 2, \cdots, 1000.$$



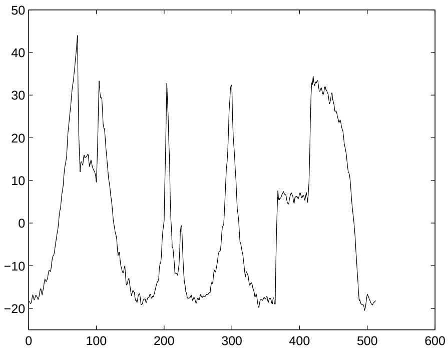
(a)



(b)



(c)



(d)

Fig. 4.1. (a) The original signal; (b) the blurred and noisy data; (c) the restored signal by the single regularization method with $\alpha = 0.08$; (d) the restored signal by the multi-parameter regularization method.

Setting $E := (e_1, e_2, \cdots, e_{1000})$, we transform the original equation (4.10) to equation

$$KEu = Ef. \tag{4.11}$$

We choose the regularization matrix Λ as a diagonal matrix with diagonal elements given by $\lambda_i = \frac{|(w,e_i)|}{|(u,e_i)|}$, $i = 1, 2, \cdots, 1000$, according to strategy 2. In Fig. 4.2, we compare the recovered signals by using the multi-parameter regularization and by the single parameter regularization, where the optimal single parameter $\alpha = 0.1$ is chosen. Again, the multi-parameter regularization performs better than the single parameter regularization.

Example 4. Consider recovering a noisy signal from a piecewise uniformly distributed noise. The original data is a piecewise constant function with a different noise magnification on different intervals. Consider solving the equation

$$f = Ku + w, \tag{4.12}$$

where f is the observed data, w is noise and K is the identity matrix.

In this experiment, we decompose the interval into three intervals and add piecewise uniform distribution noise with different intensity on the different intervals. We use the square norm of first-order derivative of u as the penalty operator. The traditional single parameter

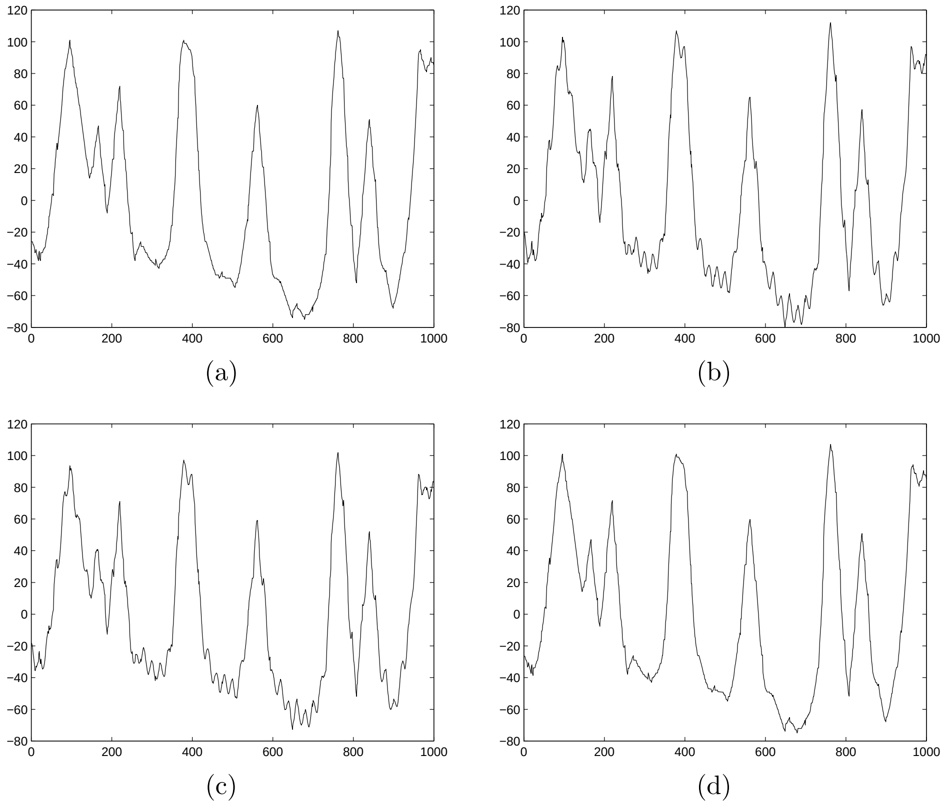


Fig. 4.2. (a) The original signal; (b) the noisy signal; (c) the recovered signal by the single parameter regularization with $\alpha = 0.1$; (d) the recovered signal by the multi-parameter regularization.

regularization method is to solve equation

$$(\alpha\Delta + K)u_\alpha = f + w, \tag{4.13}$$

where Δ is the matrix representation of the Laplace operator as a symmetric toeplitz matrix with the first row and second row as $[-2, 1, 0, \cdots, 0]$ and $[1, -2, 1, 0, \cdots, 0]$.

We use three different regularization parameters, one for each interval. Suppose that we have the decomposition

$$\mathbb{R}^{768} := \mathbb{W}_1 \oplus \mathbb{W}_2 \oplus \mathbb{W}_3,$$

where $\mathbb{W}_1 = \text{ span } \{l_i : i = 1, \cdots, 256\}$, $\mathbb{W}_2 = \text{ span } \{l_i : i = 257, \cdots, 512\}$ and $\mathbb{W}_3 = \text{ span } \{l_i : i = 513, \cdots, 768\}$, where l_i is the vector with the ith component 1 and zero elsewhere. The multi-parameter regularization method is to solve the equation

$$(\Lambda\Delta + K)u_\Lambda = f + w, \tag{4.14}$$

where $\Lambda := \text{diag}(\lambda_1 I_{256}, \lambda_2 I_{256}, \lambda_3 I_{256})$ with $\lambda_1 = 0.1, \lambda_2 = 10, \lambda_3 = 8$.

In our experiment, we use the Matlab commands "sig1 = makesignal('Riemann', 256)*10000-40; sig2 = sin((1:256)*2*2*pi/156)*100; sig3 = sin((1:256)*3*2*pi/256)*80; sig = [sig1, sig2, sig3];" to generate the original data u. We show the recovered signals in Fig. 4.3, where we compare the recovered signal by using the multi-parameter regularization with one by using

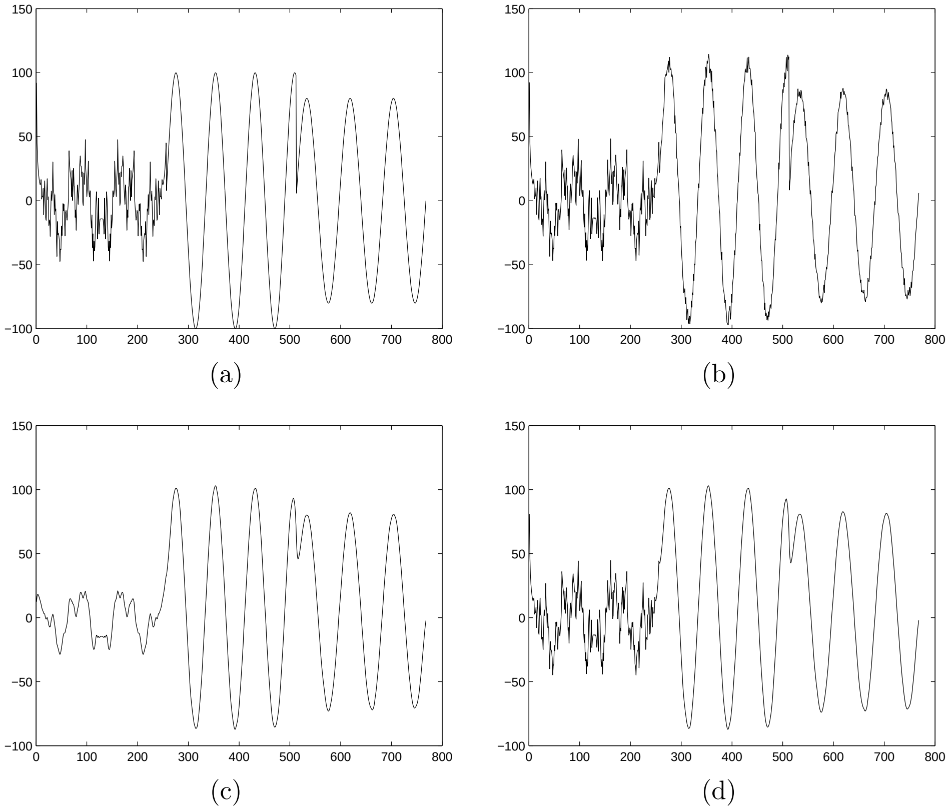


Fig. 4.3. (a) The original signal; (b) the noisy signal; (c) the recovered signal by the single parameter regularization with $\alpha = 10$; (d) the recovered signal by the multi-parameter regularization.

the single parameter regularization with the optimal parameter α. Fig. 4.3 shows that the multi-parameter regularization performs much better than the single parameter regularization.

Acknowledgments. This work was supported in part by the Natural Science Foundation of China under grants 10371137, the Foundation of Doctoral Program of National Higher Education of China under grant 20030558008, Guangdong Provincial Natural Science Foundation of China under grant 05003308 and the Foundation of Zhongshan University Advanced Research Center. It was also supported in part by the US National Science Foundation under grant CCR-0407476, by National Aeronautics and Space Administration under Cooperative Agreement NNX07AC37A, by the Natural Science Foundation of China under grants 10371122 and 10631080, and by the Education Ministry of the People's Republic of China under the Changjiang Scholar Chair Professorship Program through Zhongshan University.

References

[1] K.E. Atkinson, The Numerical Solution of Integral Equations of the Second Kind, Cambridge University Press, Cambridge, UK, 1997.

[2] M. Belge and M.E. Kilmer, E.L. Miller, Efficient determination of multiple regularization parameters in a generalized L-curve framework, *Inverse Problems*, **18** (2002), 1161-1183.

[3] C. Brezinski, M. Redivo-Zaglia, G. Rodriguez and S. Seatzu, Multi-parameter regularization techniques for ill-conditioned linear systems, *Numer. Math.*, **94** (2003), 203-228.

[4] Z. Chen, C.A. Micchelli and Y. Xu, A multilevel method for solving operator equations, *J. Math. Anal. Appl.*, **262** (2001), 688-699.

[5] Z. Chen, B. Wu, and Y. Xu, Multilevel augmentation methods for differential equations, *Adv. Comput. Math.*, **24** (2006), 213-238.

[6] Z. Chen, Y. Xu and H. Yang, A multilevel augmentation method for solving ill-posed operator equations, *Inverse Problems*, **22** (2006), 155-174.

[7] I. Daubechies, *Ten Lectures on Wavelets*, CBMS-NSF Regional Conference Series in Appl. Math. **61**, SIAM, Philadelphia, 1992.

[8] A. Doicu, F. Schreier, S. Hilgers and M. Hess, Multi-parameter regularization method for atmospheric remote sensing, *Comput. Phys. Commun.*, **165** (2005), 1-9.

[9] H.W. Engl, M. Hanke and A. Neubauer, *Regularization of Inverse Problems*, Kluwer, Dordrecht, 1996.

[10] W. Fang and M. Lu, A fast collocation method for an inverse boundary value problem, *Int. J. Numer. Methods Engrg.*, **59** (2004), 1563-1585.

[11] H. Gfrerer, An a posteriori parameter choice for ordinary and iterated Tikhonov regularization of ill-posed problems leading to optimal convergence rates, *Math. Comput.*, **49** (1987), 507-522.

[12] C.W. Groetsch, *The Theory of Tikhonov Regularization for Fredholm Equations of the First Kind*, Research Notes in Mathematics, 105, Pitman, Boston, MA, 1984.

[13] Y. Lu, L. Shen and Y. Xu, Shadow block iteration for solving linear systems obtained from wavelet transforms, *Appl. Comput. Harmonic Anal.*, **19** (2005), 259-385.

[14] Y. Lu, L. Shen, and Y. Xu, Multi-parameter regularization methods for high-resolution image reconstruction with displacement errors, *IEEE T. Circuits–I*, **54** (2007), 1788-1799.

[15] K. Miller, Least squares methods for ill-posed problems with a prescribed bound, *SIAM J. Math. Anal.*, **1** (1970), 52-74.

[16] C.A. Micchelli and Y. Xu, Using the matrix refinement equation for the construction of wavelets on invariant sets, *Appl. Comput. Harmonic Anal.*, **1** (1994), 391-401.

[17] C.A. Micchelli and Y. Xu, Reconstruction and decomposition algorithms for biorthogonal multiwavelets, *Multidim. Syst. Sign. Process.*, **8** (1997), 31-69.

[18] D.L. Phillips, A technique for the numerical solution of certain integral equations of the first kind, *J. Assoc. Comput. Mach.*, **9** (1962), 84-97.

[19] R. Plato, *Iterative and parametric methods for linear ill-posed equations*, Habilitationsschrift, Fachbereich Mathematik, TU Berlin, 1995.

[20] T. Raus, The principle of the residual in the solution of ill-posed problems with nonselfadjoint operator (Russian), *Acta et Comment. Univ. Tartuensis*, **715** (1985), 12-20.

[21] U. Tautenhahn and U. Hämarik, The use of monotonicity for choosing the regularization parameter in ill-posed problems, *Inverse Problems*, **15** (1999), 1487-1505.

[22] A.N. Tikhonov, On the solution of ill-posed problems and the method of regularization (Russian), *Dokl. Acad. Nauk USSR*, **151** (1963), 501-504.

[23] A.N. Tikhonov and V.Y. Arsenin, *Solutions of Ill-posed Problems*, Wiley, New York, 1977.

[24] G. Vainikko, The discrepancy principle for a class of regularization methods, *USSR Comp. Math. Phys.*, **22** (1982), 1-19.

[25] J. Xie and J. Zou, An improved model function method for choosing regularization parameters in linear inverse problems, *Inverse Problems*, **18** (2002), 631-643.