Doctoral Program
Computational Mathematics
Numerical Analysis and Symbolic Computation

A Double Regularization Approach for Inverse Problems with Noisy Data and Inexact Operator

Ismael Rodrigo Bleyer Ronny Ramlau

DK-Report No. 2012-15 12 2012

A–4040 LINZ, ALTENBERGERSTRASSE 69, AUSTRIA

Supported by
Austrian Science Fund (FWF) Upper Austria

DK sponsors:

- **Johannes Kepler University Linz** (JKU)
- **Austrian Science Fund** (FWF)
- **Upper Austria**

A Double Regularization Approach for Inverse Problems with Noisy Data and Inexact Operator

Ismael Rodrigo Bleyer[†] Ronny Ramlau[‡]

November 13, 2012

Abstract

In standard inverse problems, the task is to solve an operator equation from given noisy data. However, sometimes also the operator is not known exactly. Therefore we propose a method that allows errors both in the operator and the data. In particular, we consider operator equations where the operator can be characterized by a function. For the stable reconstruction we propose the use of a Tikhonov-type functional with a generalized misfit term and an additional penalty term which promotes sparsity. Using an appropriate parameter choice rule for the two regularization parameters, we prove convergence and convergence rates for the method, and provide a first numerical example.

1 Introduction

In this paper, we consider the inversion of a linear operator $A_0 : \mathcal{V} \to \mathcal{H}$ defined between Hilbert spaces, i.e., we want to solve the equation

$$A_0 f = g_0 \,. \tag{1}$$

Additionally we will assume that only noisy data g_δ with

$$\|g_0 - g_\delta\|_{\mathcal{H}} \le \delta$$

is available. If the problem of solving (1) is *ill-posed*, i.e., the solution of the equation depends not continuously on the data [12], then even small noise in the data can create large deviations in the reconstructions. In this case, regularization methods have to be used for a stable inversion. In standard Inverse Problems theory it is assumed that the operator A_0 is known exactly. This, however, is not always the case, as in some applications only

[†]ismael.bleyer@dk-compmath.jku.at
[‡]ronny.ramlau@jku.at

an approximation A_ϵ of the exact operator is known. A typical example from imaging is a deconvolution problem with approximately known or unknown convolution kernel, as, e.g., it was the case for early Hubble images [3, 14, 4, 1]. Another example is connected to Inverse Scattering, where the Linear Sampling Method involves the solution of an integral equation with approximately known kernel, see [5] and references therein.

Over the last decades, several approaches have been proposed that consider the inversion of an equation with both noise in the data and in the operator. Most of the papers published in journals focus on the finite dimensional setup. However, an entire chapter in the book [24] is devoted to solving the problem in Hilbert spaces. Unfortunately the book is only available in Russian and thus is not easily accessible.

In 1980, Golub and Van Loan [11] investigated a fitting technique based on the least squares problem for solving a matrix equation with incorrect matrix and data vector, the so-called *total least squares* (TLS) method, for more details see [25, 19]. Later, Tikhonov regularization was recast as a TLS formulation resulting in the *regularized total least squares* method (R-TLS), see [11, 13, 10].

In a finite dimensional setting[1], the R-TLS method can be formulated as constrained minimization problem:

$$\begin{aligned} &\text{minimize} && \left\|A - A_\epsilon\right\|_F^2 + \left\|g - g_\delta\right\|_2^2 \\ &\text{subject to} && \begin{cases} Af = g \\ \left\|Lf\right\|_2^2 \le M\,. \end{cases} \end{aligned}$$

The optimal pair (A, g) minimizes the residual in the operator and in the data, measured by Frobenius and Euclidian norm, respectively. Moreover, the solution pair is connected via the equation $Af = g$, where the element f belongs to a ball in $\mathcal{V}$ of radius M. The "size" of the ball is measured by a linear and invertible operator L (often the identity). Any element f satisfying these constraints defines a R-TLS solution.

The accuracy of the R-TLS depends heavily on the right choice of M, which is usually difficult to obtain. An alternative is the *dual regularized total least square* (D-RTLS) method, where the approximation f to the solution of the equation (1) is given as the minimizer of the following problem

$$\begin{aligned} &\text{minimize} && \left\|Lf\right\|_2^2 \\ &\text{subject to} && \begin{cases} Af = g \\ \left\|g - g_\delta\right\|_2^2 \le \delta \\ \left\|A - A_\epsilon\right\|_F^2 \le \epsilon\,, \end{cases} \end{aligned}$$

where $\left\|\cdot\right\|_F$ denotes again the Frobenius norm. Please note that most of the

[1] we keep the same notation as in the infinite dimensional setup.

available results on this method do again require a finite dimensional setup, see, e.g., [17, 18, 23].

In our approach, we would like to restrict our attention to linear operators that can be mainly characterized by a function, as it is, e.g., the case for linear integral operators, where the kernel function determines the behavior of the operator. Moreover, we will assume that the noise in the operator is due to an incorrect characterizing function. This approach will allow us to treat the problem of finding a solution of (1) from incorrect data and operator in the framework of Tikhonov regularization rather than as a constraint minimization problem.

The further contents of the paper is organized as follows: in Section 2 we formulate the underlying problem and we include few examples as motivation. In Section 3 we introduce the proposed method as well as its mathematical setting. We analyze its regularization properties: existence, stability and convergence in Section 4. Additionally we study source condition and derive convergence rates with respect to Bregman distance. Finally, in Section 5, we shortly comment computational issues through a numerical illustration.

2 Problem formulation and examples

As mentioned above, we aim at the inversion of linear operator equation $A_0 f = g_0$ from noisy data g_δ and incorrect operator A_ϵ. Additionally we assume that the operators $A_0, A_\epsilon : \mathcal{V} \to \mathcal{H}$, $\mathcal{V}, \mathcal{H}$ Hilbert spaces, can be characterized by functions $k_0, k_\epsilon \in \mathcal{U}$. To be more specific, we consider operators

$$\begin{aligned} A_k : \quad \mathcal{V} &\longrightarrow \quad \mathcal{H} \\ v &\longmapsto \quad B(k,v) \ , \end{aligned}$$

i.e., $A_k v = B(k,v)$, where B is a *bilinear* operator

$$B : \mathcal{U} \times \mathcal{V} \to \mathcal{H}$$

fulfilling, for some $C > 0$,

$$\|B(k,f)\|_{\mathcal{H}} \le C \|k\|_{\mathcal{U}} \|f\|_{\mathcal{V}}. \tag{2}$$

From (2) follows immediately

$$\|B(k,\cdot)\|_{\mathcal{V}\to\mathcal{H}} \le C \|k\|_{\mathcal{U}}. \tag{3}$$

Associated with the bilinear operator B, we also define the linear operator

$$\begin{aligned} C_f : \quad \mathcal{U} &\longrightarrow \quad \mathcal{H} \\ u &\longmapsto \quad B(u,f) \ , \end{aligned}$$

i.e., $C_f u = B(u, f)$.
From now on, let us identify A_0 with A_{k_0} and A_ϵ with A_{k_ϵ}. From (3) we deduce immediately

$$\|A_0 - A_\epsilon\| \leq C\|k_0 - k_\epsilon\| \ , \tag{4}$$

i.e., the operator error norm is controlled by the error norm of the characterizing functions. Now we can formulate our problem as follows:

$$\begin{aligned} \text{Solve} \qquad & A_0 f = g_0 && \text{(5a)} \\ \text{from noisy data } g_\delta \text{ with} \qquad & \|g_0 - g_\delta\| \leq \delta && \text{(5b)} \\ \text{and noisy function } k_\epsilon \text{ with} \qquad & \|k_0 - k_\epsilon\| \leq \epsilon \ . && \text{(5c)} \end{aligned}$$

Please note that the problem with explicitly known k_0 (or the operator A_0) is often ill-posed and needs regularization for a stable inversion. Therefore we will also propose a regularizing scheme for the problem (5a)-(5c). Now let us give some examples.

Example 1. Consider a linear integral operator A_0 defined through

$$(A_0 f)(s) := \int_\Omega k_0(s,t) f(t) dt = B(k_0, f)$$

with $\mathcal{V} = \mathcal{H} = L_2(\Omega)$ and let k_0 be a function in $\mathcal{U} = L_2(\Omega^2)$. Then the bilinear operator B yields

$$\|B(k_0, f)\| \leq \|k_0\|_{\mathcal{U}} \, \|f\|_{\mathcal{V}}.$$

The considered class of operators also contains deconvolution problems, which are important in imaging, as well as *blind deconvolution* problems [15, 3, 14], where it is assumed that also the exact convolution kernel is unknown.

Example 2. In medical imaging, the data of *Single Photon Emission Computed Tomography* (SPECT) is described by the attenuated Radon transform [20, 7, 21]:

$$Af(s,\omega) = \int_{\mathbb{R}} f(s\omega^\perp + t\omega) \cdot e^{-\int_t^\infty \mu(s\omega^\perp + \tau\omega)\, d\tau} dt \ .$$

The function μ is the density distribution of the body. In general, the density distribution is also unknown. Modern scanner, however, perform a CT scan in parallel. Due to measurement errors, the reconstructed density distribution is also incorrect. Setting

$$k_\epsilon(s,t,\omega) = e^{-\int_t^\infty \mu_\epsilon(s\omega^\perp + \tau\omega)\, d\tau} \ ,$$

we have

$$A_\epsilon f = B(k_\epsilon, f) \ ,$$

and similar estimates as in (2) can be obtained.

3 Double Regularized Total Least Squares (dbl-RTLS)

Due to our assumptions on the structure of the operator A_0, the inverse problem of identifying the function f^{true} from noisy measurements g_δ and inexact operator A_ϵ can now be rewritten as the task of solving the inverse problem

$$B(k_0, f) = g_0 \tag{6}$$

from noisy measurements (k_ϵ, g_δ) fulfilling

$$\|g_0 - g_\delta\|_{\mathcal{H}} \leq \delta, \tag{7a}$$

and

$$\|k_0 - k_\epsilon\|_{\mathcal{U}} \leq \epsilon. \tag{7b}$$

In most applications, the "inversion" of B will be ill-posed (e.g., if B is defined via a Fredholm integral operator), and a regularization strategy is needed for a stable solution of the problem (6).

The structure of our problem allows to reformulate (5a)-(5c) as an unconstrained Tikhonov-type problem:

$$\underset{(k,f)}{\text{minimize}} \ J^{\delta,\varepsilon}_{\alpha,\beta}\big(k, f\big) := \frac{1}{2} T^{\delta,\varepsilon}\left(k, f\right) + R_{\alpha,\beta}\left(k, f\right) \ , \tag{8a}$$

where

$$T^{\delta,\varepsilon}\left(k, f\right) = \big\|B(k, f) - g_\delta\big\|^2 + \gamma\big\|k - k_\epsilon\big\|^2 \tag{8b}$$

and

$$R_{\alpha,\beta}\left(k, f\right) = \frac{\alpha}{2}\big\|Lf\big\|^2 + \beta\mathcal{R}(k). \tag{8c}$$

Here, α and β are the regularization parameters which have to be chosen properly, γ is a scaling parameter, L is a bounded linear and continuously invertible operator and $\mathcal{R} : X \subset \mathcal{U} \to [0, +\infty]$ is proper, convex and weakly lower semi-continuous functional . We wish to note that most of the available papers assume that L is a densely defined, unbounded self-adjoint and strictly positive operator, see, e.g. [17, 16]. For our analysis, however, boundedness is needed and it is an open question whether the analysis could be extended to cover unbounded operators, too.

We call this scheme the *double regularized total least squares method* (dbl-RTLS). Please note that the method is closely related to the total least squares method, as the term $\big\|k - k_\epsilon\big\|^2$ controls the error in the operator. The functional $J_{\alpha,\beta}^{\delta,\varepsilon}$ is composed as the sum of two terms: one which measures the discrepancy of data and operator, and one which promotes stability. The functional $T^{\delta,\varepsilon}$ is a *data-fidelity* term based on the TLS technique, whereas the functional $R_{\alpha,\beta}$ acts as a *penalty* term which stabilizes the inversion with respect to the pair (k, f). As a consequence, we have two regularization parameters, which also occurs in *double regularization*, see, e.g., [26].

The domain of the functional $J_{\alpha,\beta}^{\delta,\varepsilon} : \big(\mathcal{U} \cap X\big) \times \mathcal{V} \longrightarrow \mathbb{R}$ can be extended over $\mathcal{U} \times \mathcal{V}$ by setting $\mathcal{R}(k) = +\infty$ whenever $k \in \mathcal{U} \setminus X$. Then $\mathcal{R}$ is proper, convex and weak lower semi-continuous functional in $\mathcal{U}$.

4 Well-posedness and convergence rates

In this Section we shall analyze some analytical properties of the proposed dbl-RTLS method. In particular, we prove its well-posedness as a regularization method, i.e., the minimizers of the regularization functional $J_{\alpha,\beta}^{\delta,\varepsilon}$ exist for every $\alpha, \beta > 0$, depend continuously on both g_δ and k_ϵ, and converge to a solution of $B(k_0, f) = g_0$ as both noise level approaches zero, provided the regularization parameters α and β are chosen appropriately.

For the pair $(k, f) \in \mathcal{U} \times \mathcal{V}$ we use the canonical inner product

$$\langle (k_1, f_1), (k_2, f_2) \rangle_{\mathcal{U}\times\mathcal{V}} := \langle k_1, k_2 \rangle_{\mathcal{U}} + \langle f_1, f_2 \rangle_{\mathcal{V}} \ ,$$

i.e., convergence is defined componentwise. For the upcoming results, we need the following assumption on the operator B:

Assumption A.

(A1) B is strongly continuous, i.e., if $(k^n, f^n) \rightharpoonup (\bar{k}, \bar{f})$ then $B(k^n, f^n) \to B(\bar{k}, \bar{f})$.

Proposition 4.1. *Let $J_{\alpha,\beta}^{\delta,\varepsilon}$ be the functional defined in (8). Assume that L is a bounded linear and continuously invertible operator and B fulfills Assumption A1. Then $J_{\alpha,\beta}^{\delta,\varepsilon}$ is a positive, weakly lower semi-continuous and coercive functional.*

Proof. By the definition of $T^{\delta,\varepsilon}$, $\mathcal{R}$ and Assumption A1, $J_{\alpha,\beta}^{\delta,\varepsilon}$ is positive and w-lsc. As the operator L is continuously invertible, there exists a constant $c > 0$ such that

$$c\big\|f\big\| \le \big\|Lf\big\|$$

for all $f \in \mathscr{D}(L)$. We get

$$J_{\alpha,\beta}^{\delta,\varepsilon}\big(k, f\big) \ge \gamma\|k - k_\epsilon\|^2 + \frac{\alpha c}{2}\|f\|^2 \to \infty$$

as $\big\|(k, f)\big\|^2 := \big\|k\big\|^2 + \big\|f\big\|^2 \to \infty$ and therefore $J_{\alpha,\beta}^{\delta,\varepsilon}$ is coercive. ∎

We point out here that the problem (6) may not even have a solution for any given noisy measurements (k_ϵ, g_δ) whereas the regularized problem (8) does, as stated below:

Theorem 4.2 (Existence)**.** *Let the assumptions of Proposition 4.1 hold. Then the functional* $J^{\delta,\varepsilon}_{\alpha,\beta}\big(k, f\big)$ *has a global minimizer.*

Proof. By Proposition 4.1, $J^{\delta,\varepsilon}_{\alpha,\beta}\big(k, f\big)$ is positive, proper and coercive, i.e., there exists $(k, f) \in \mathscr{D}(J^{\delta,\varepsilon}_{\alpha,\beta})$ such that $J^{\delta,\varepsilon}_{\alpha,\beta}\big(k, f\big) < \infty$.

Let $\nu = \inf\{\, J^{\delta,\varepsilon}_{\alpha,\beta}\big(k, f\big) \mid (k, f) \in \operatorname{dom} J^{\delta,\varepsilon}_{\alpha,\beta}\}$. Then, there exists $M > 0$ and a sequence $(k^j, f^j) \in \operatorname{dom} J^{\delta,\varepsilon}_{\alpha,\beta}$ such that $J(k^j, f^j) \to \nu$ and

$$J^{\delta,\varepsilon}_{\alpha,\beta}\big(k^j, f^j\big) \le M \quad \forall j.$$

In particular we have

$$\frac{1}{2}\alpha\big\|Lf^j\big\|^2 \le M \quad \text{and} \quad \frac{1}{2}\gamma\big\|k^j - k_\epsilon\big\|^2 \le M.$$

Using

$$\big\|k^j\big\| - \big\|k_\epsilon\big\| \le \big\|k^j - k_\epsilon\big\| \le \left(\frac{2M}{\gamma}\right)^{1/2}$$

it follows

$$\big\|k^j\big\| \le \left(\frac{2M}{\gamma}\right)^{1/2} + \big\|k_\epsilon\big\| \quad \text{and} \quad \big\|f^j\big\| \le \left(\frac{2M}{\alpha c^2}\right)^{1/2},$$

i.e., the sequences (k^j) and (f^j) are bounded. Thus there exist subsequences of (k^j), (f^j) (for simplicity, again denoted by (k^j) and (f^j)) s.t.

$$k^j \rightharpoonup \bar{k} \quad \text{and} \quad f^j \rightharpoonup \bar{f},$$

and thus

$$(k^j, f^j) \rightharpoonup (\bar{k}, \bar{f}) \in (\mathcal{U} \cap X) \times \mathcal{V}.$$

By the w-lsc of the functional $J^{\delta,\varepsilon}_{\alpha,\beta}$ we obtain

$$\nu \le J^{\delta,\varepsilon}_{\alpha,\beta}\big(\bar{k}, \bar{f}\big) \le \liminf J^{\delta,\varepsilon}_{\alpha,\beta}\big(k^j, f^j\big) = \lim J^{\delta,\varepsilon}_{\alpha,\beta}\big(k^j, f^j\big) = \nu$$

Hence $\nu = J^{\delta,\varepsilon}_{\alpha,\beta}\big(\bar{k}, \bar{f}\big)$ is the minimum of the functional and $(\bar{k}, \bar{f})$ is a global minimizer,

$$(\bar{k}, \bar{f}) = \arg\min\{\, J^{\delta,\varepsilon}_{\alpha,\beta}\big(k, f\big) \mid (k, f) \in \mathscr{D}(J^{\delta,\varepsilon}_{\alpha,\beta})\}.$$

□

The stability property of the standard Tikhonov regularization strategy for problems with noisy right hand side is well known. We next investigate this property for the Tikhonov-type regularization scheme (8) for perturbations on both (k_ϵ, g_δ).

Theorem 4.3 (Stability)**.** *Let $\alpha, \beta > 0$ be fixed the regularization parameters, L a bounded and continuously invertible operator and $(g_{\delta_j})_j$, $(k_{\epsilon_j})_j$ sequences with $g_{\delta_j} \to g_\delta$ and $k_{\epsilon_j} \to k_\epsilon$. If (k^j, f^j) denote minimizers of $J_{\alpha,\beta}^{\delta_j,\varepsilon_j}$ with data g_{δ_j} and characterizing function k_{ϵ_j}, then there exists a convergent subsequence of $(k^j, f^j)_j$. The limit of every convergent subsequence is a minimizer of the functional $J_{\alpha,\beta}^{\delta,\varepsilon}$.*

Proof. By the definition of (k^j, f^j) as minimizers of $J_{\alpha,\beta}^{\delta_j,\varepsilon_j}$ we have

$$J_{\alpha,\beta}^{\delta_j,\varepsilon_j}\left(k^j, f^j\right) \le J_{\alpha,\beta}^{\delta_j,\varepsilon_j}\left(k, f\right) \quad \forall (k,f) \in \mathscr{D}(J_{\alpha,\beta}^{\delta,\varepsilon}), \tag{9}$$

With $(\tilde{k}, \tilde{f}) := (k_{\alpha,\beta}^{\delta,\epsilon}, f_{\alpha,\beta}^{\delta,\epsilon})$ we get $J_{\alpha,\beta}^{\delta_j,\varepsilon_j}\left(\tilde{k}, \tilde{f}\right) \to J_{\alpha,\beta}^{\delta,\varepsilon}\left(\tilde{k}, \tilde{f}\right)$. Hence, there exists a $\tilde{c} > 0$ so that $J_{\alpha,\beta}^{\delta_j,\varepsilon_j}\left(\tilde{k}, \tilde{f}\right) \le \tilde{c}$ for j sufficiently large. In particular, we observe with (9) that $\left(\left\|k^j - k_{\epsilon_j}\right\|\right)_j$ as well as $\left(\left\|Lf^j\right\|\right)_j$ are uniformly bounded.

Analogous to the proof of Theorem 4.2 we conclude that the sequence $(k^j, f^j)_j$ is uniformly bounded. Hence there exists a subsequence (for simplicity also denoted by$(k^j, f^j)_j$) such that

$$k^j \rightharpoonup \bar{k} \quad \text{and} \quad f^j \rightharpoonup \bar{f}.$$

By the weak lower semicontinuity (w-lsc) of the norm and continuity of B we have

$$\left\|B(\bar{k}, \bar{f}) - g_\delta\right\| \le \liminf_j \left\|B(k^j, f^j) - g_{\delta_j}\right\|$$

and

$$\left\|\bar{k} - k_\epsilon\right\| \le \liminf_j \left\|k^j - k_{\epsilon_j}\right\|.$$

Moreover, (9) implies

$$\begin{aligned} J_{\alpha,\beta}^{\delta,\varepsilon}\left(\bar{k}, \bar{f}\right) &\le \liminf_j J_{\alpha,\beta}^{\delta_j,\varepsilon_j}\left(k^j, f^j\right) \\ &\le \limsup_j J_{\alpha,\beta}^{\delta_j,\varepsilon_j}\left(k, f\right) \\ &= \lim_j J_{\alpha,\beta}^{\delta_j,\varepsilon_j}\left(k, f\right) \\ &= J_{\alpha,\beta}^{\delta,\varepsilon}\left(k, f\right) \end{aligned}$$

for all $(k, f) \in \mathscr{D}(J_{\alpha,\beta}^{\delta,\varepsilon})$. In particular, $J_{\alpha,\beta}^{\delta,\varepsilon}\left(\bar{k}, \bar{f}\right) \le J_{\alpha,\beta}^{\delta,\varepsilon}\left(\tilde{k}, \tilde{f}\right)$. Since $(\tilde{k}, \tilde{f})$ is by definition a minimizer of $J_{\alpha,\beta}^{\delta,\varepsilon}$, we conclude $J_{\alpha,\beta}^{\delta,\varepsilon}\left(\bar{k}, \bar{f}\right) = J_{\alpha,\beta}^{\delta,\varepsilon}\left(\tilde{k}, \tilde{f}\right)$ and thus

$$\lim_{j\to\infty} J_{\alpha,\beta}^{\delta_j,\varepsilon_j}\left(k^j, f^j\right) = J_{\alpha,\beta}^{\delta,\varepsilon}\left(\bar{k}, \bar{f}\right). \tag{10}$$

It remains to show

$$k^j \to \bar{k} \quad \text{and} \quad f^j \to \bar{f}.$$

As the sequences are weakly convergent, convergence of the sequences holds if

$$\|k^j\| \to \|\bar{k}\| \quad \text{and} \quad \|f^j\| \to \|\bar{f}\|.$$

The norms on $\mathcal{U}$ and $\mathcal{V}$ are w-lsc, thus it is sufficient to show

$$\|\bar{k}\| \geq \limsup\|k^j\| \quad \text{and} \quad \|\bar{f}\| \geq \limsup\|f^j\|.$$

The operator L is bounded and continuously invertible, therefore $f^j \to \bar{f}$ if and only if $Lf^j \to L\bar{f}$. Therefore, we accomplish the prove for the sequence $(Lf^j)_j$. Now suppose there exists τ_1 as

$$\tau_1 := \limsup\|Lf^j\| > \|L\bar{f}\|$$

and there exists a subsequence $(f^n)_n$ of $(f^j)_j$ such that $Lf^n \rightharpoonup L\bar{f}$ and $\|Lf^n\| \to \tau_1$.
From the first part of this proof (10), it holds

$$\lim_{j\to\infty} J_{\alpha,\beta}^{\delta_j,\varepsilon_j}\left(k^j, f^j\right) = J_{\alpha,\beta}^{\delta,\varepsilon}\left(\bar{k}, \bar{f}\right).$$

Using (8) we observe

$$\begin{aligned}
&\lim_{n\to\infty}\left(\frac{1}{2}\|B(k^n, f^n) - g_{\delta_n}\|^2 + \frac{\gamma}{2}\|k^n - k_{\epsilon_n}\|^2 + \beta\mathcal{R}(k^n)\right)\\
= \;& \frac{1}{2}\|B(\bar{k}, \bar{f}) - g_\delta\|^2 + \frac{\gamma}{2}\|\bar{k} - k_\epsilon\|^2 + \beta\mathcal{R}(\bar{k}) + \frac{\alpha}{2}\left(\|L\bar{f}\|^2 - \lim_{n\to\infty}\|Lf^n\|^2\right)\\
= \;& \frac{1}{2}\|B(\bar{k}, \bar{f}) - g_\delta\|^2 + \frac{\gamma}{2}\|\bar{k} - k_\epsilon\|^2 + \beta\mathcal{R}(\bar{k}) + \frac{\alpha}{2}\left(\|L\bar{f}\|^2 - {\tau_1}^2\right)\\
< \;& \frac{1}{2}\|B(\bar{k}, \bar{f}) - g_\delta\|^2 + \frac{\gamma}{2}\|\bar{k} - k_\epsilon\|^2 + \beta\mathcal{R}(\bar{k})\;,
\end{aligned}$$

which is a contradiction to the w-lsc property of the involved norms and the functional $\mathcal{R}$. Thus $Lf^j \to L\bar{f}$ and

$$f^j \to \bar{f}.$$

The same idea can be used in order to prove convergence of the characterizing functions. Suppose there exists τ_2 s.t.

$$\tau_2 := \limsup\|k^j - k_\epsilon\| > \|\bar{k} - k_\epsilon\|$$

and there exists a subsequence $(k^n)_n$ of $(k^j)_j$ such that $(k^n - k_\epsilon) \rightharpoonup (\bar{k} - k_\epsilon)$ and $\|k^n - k_\epsilon\| \to \tau_2$.

By the triangle inequality we get

$$\|k^n - k_\epsilon\| - \|k_{\epsilon_n} - k_\epsilon\| \leq \|k^n - k_{\epsilon_n}\| \leq \|k^n - k_\epsilon\| + \|k_{\epsilon_n} - k_\epsilon\|,$$

and thus

$$\lim_{n\to\infty}\left\|k^n - k_{\epsilon_n}\right\| = \lim_{n\to\infty}\left\|k^n - k_\epsilon\right\| .$$

Therefore

$$\begin{aligned}
& \lim_{n\to\infty}\left(\frac{1}{2}\left\|B(k^n, f^n) - g_{\delta_n}\right\|^2 + \beta\mathcal{R}(k^n)\right) \\
= \;& \frac{1}{2}\left\|B(\bar{k}, \bar{f}) - g_\delta\right\|^2 + \frac{\gamma}{2}\left(\left\|\bar{k} - k_\epsilon\right\|^2 - \lim_{n\to\infty}\left\|k^n - k_{\epsilon_n}\right\|^2\right) + \beta\mathcal{R}(\bar{k}) \\
= \;& \frac{1}{2}\left\|B(\bar{k}, \bar{f}) - g_\delta\right\|^2 + \frac{\gamma}{2}\left(\left\|\bar{k} - k_\epsilon\right\|^2 - {\tau_2}^2\right) + \beta\mathcal{R}(\bar{k}) \\
< \;& \frac{1}{2}\left\|B(\bar{k}, \bar{f}) - g_\delta\right\|^2 + \beta\mathcal{R}(\bar{k}) \ ,
\end{aligned}$$

which is again a contradiction to the w-lsc of the involved norms and functionals. □

4.1 Convergence

In the following, we investigate the regularization property of our approach, i.e., we show, under an appropriate parameter choice rule, that the minimizers $(k^{\delta,\epsilon}_{\alpha,\beta}, f^{\delta,\epsilon}_{\alpha,\beta})$ of the functional (8) converge to an exact solution as the noise level (δ, ϵ) goes to zero.

Let us first clarify our notion of a solution. In principle, the equation $B(k, f) = g$ might have different pairs (k, f) as solution. However, as $k_\epsilon \to k_0$ as $\epsilon \to 0$, we get k_0 for free in the limit, that is, we are interested in reconstructing solutions of the equation $B(k_0, f) = g$. In particular, we want to reconstruct a solution with minimal value of $\|Lf\|$, and therefore define:

Definition 4.4. *We call $f^\dagger$ a minimum-norm solution if*

$$f^\dagger = \arg\min_f \left\{\|Lf\| \;\mid\; B(k_0, f) = g_0\right\} .$$

The definition above is the standard *minimum-norm solution* for the classical Tikhonov regularization (see for instance [9]).

Furthermore, we have to introduce a regularization parameter choice which depends on both noise level, defined through (11) in the upcoming theorem.

Theorem 4.5 (convergence)**.** *Let the sequences of data g_{δ_j} and k_{ϵ_j} with $\left\|g_{\delta_j} - g_0\right\| \le \delta_j$ and $\left\|k_{\epsilon_j} - k_0\right\| \le \epsilon_j$ be given with $\epsilon_j \to 0$ and $\delta_j \to 0$. Assume that the regularization parameters $\alpha_j = \alpha(\epsilon_j, \delta_j)$ and $\beta_j = \beta(\epsilon_j, \delta_j)$ fulfill $\alpha_j \to 0$, $\beta_j \to 0$, as well as*

$$\lim_{j\to\infty}\frac{\delta_j^2 + \gamma\epsilon_j^2}{\alpha_j} = 0 \quad \textit{and} \quad \lim_{j\to\infty}\frac{\beta_j}{\alpha_j} = \eta \tag{11}$$

for some $0 < \eta < \infty$.

Let the sequence

$$(k^j, f^j)_j := \left(k_{\alpha_j,\beta_j}^{\delta_j,\epsilon_j}, f_{\alpha_j,\beta_j}^{\delta_j,\epsilon_j}\right)_j$$

be the minimizer of (8)*, obtained from the noisy data* g_{δ_j} *and* k_{ϵ_j}*, regularization parameters* α_j *and* β_j *and scaling parameter* γ.

Then there exists a convergent subsequence of $(k^j, f^j)_j$ *with* $k^j \to k_0$ *and the limit of every convergent subsequence of* $(f^j)_j$ *is a minimum-norm solution of* (6).

Proof. The minimizing property of (k^j, f^j) guarantees

$$J_{\alpha_j,\beta_j}^{\delta_j,\varepsilon_j}(k^j, f^j) \le J_{\alpha_j,\beta_j}^{\delta_j,\varepsilon_j}(k, f), \quad \forall (k, f) \in \mathscr{D}(J_{\alpha,\beta}^{\delta,\varepsilon}).$$

In particular,

$$0 \le J_{\alpha_j,\beta_j}^{\delta_j,\varepsilon_j}(k^j, f^j) \le J_{\alpha_j,\beta_j}^{\delta_j,\varepsilon_j}(k_0, f^\dagger) \le \frac{\delta_j^2 + \gamma\epsilon_j^2}{2} + \frac{\alpha_j}{2}\left\|Lf^\dagger\right\|^2 + \beta_j \mathcal{R}(k_0), \quad (12)$$

where $f^\dagger$ denotes a minimum-norm solution of the equation $B(k_0, f) = g_0$, see Definition 4.4.

Combining this estimate with the assumptions on the regularization parameters, we conclude that the sequences

$$\left\|B(k^j, f^j) - g_{\delta_j}\right\|^2, \left\|k^j - k_{\epsilon_j}\right\|^2, \left\|Lf^j\right\|^2, \mathcal{R}(k^j)$$

are uniformly bounded and by the invertibility of L, the sequence $(k^j, f^j)_j$ is uniformly bounded.

Therefore it exists a weakly convergent subsequence $(k^m, f^m)_m := (k^{j_m}, f^{j_m})_{j_m}$ of $(k^j, f^j)_j$ with

$$(k^m, f^m) \rightharpoonup (\bar{k}, \bar{f}).$$

In the following we will prove that for the weak limit $(\bar{k}, \bar{f})$ holds $\bar{k} = k_0$ and $\bar{f}$ is a minimum-norm solution.

By the weak lower semi-continuity of the norm we have

$$\begin{aligned}
0 &\le \frac{1}{2}\left\|B(\bar{k}, \bar{f}) - g_0\right\|^2 + \frac{\gamma}{2}\left\|\bar{k} - k_0\right\|^2 \\
&\le \liminf_{m\to\infty} \left\{\frac{1}{2}\left\|B(k^m, f^m) - g_{\delta_m}\right\|^2 + \frac{\gamma}{2}\left\|k^m - k_{\epsilon_m}\right\|^2\right\} \\
&\overset{(12)}{\le} \liminf_{m\to\infty} \left\{\frac{\delta_m^2 + \gamma\epsilon_m^2}{2} + \frac{\alpha_m}{2}\left\|Lf^\dagger\right\|^2 + \beta_m \mathcal{R}(k_0)\right\} \\
&= 0,
\end{aligned}$$

where the last equality follows from the parameter choice rule.

In particular, we have

$$\bar{k} = k_0 \quad \text{and} \quad B(\bar{k}, \bar{f}) = g_0.$$

From (12) follows

$$\frac{1}{2}\left\|Lf^m\right\|^2 + \frac{\beta_m}{\alpha_m}\mathcal{R}(k^m) \leq \frac{\delta_m^2 + \gamma\epsilon_m^2}{2\alpha_m} + \frac{1}{2}\left\|Lf^\dagger\right\|^2 + \frac{\beta_m}{\alpha_m}\mathcal{R}(k_0)\ .$$

Again, weak lower semi-continuity of the norm and the functional $\mathcal{R}$ result in

$$\begin{aligned}
\frac{1}{2}\left\|L\bar{f}\right\|^2 + \eta\mathcal{R}(\bar{k}) &\leq \liminf_{m\to\infty}\left\{\frac{1}{2}\left\|Lf^m\right\|^2 + \eta\mathcal{R}(k^m)\right\} \\
&= \liminf_{m\to\infty}\left\{\frac{1}{2}\left\|Lf^m\right\|^2 + \frac{\beta_m}{\alpha_m}\mathcal{R}(k^m)\right\} \\
&\leq \liminf_{m\to\infty}\left\{\frac{\delta_m^2 + \gamma\epsilon_m^2}{2\alpha_m} + \frac{1}{2}\left\|Lf^\dagger\right\|^2 + \frac{\beta_m}{\alpha_m}\mathcal{R}(k_0)\right\} \\
&\overset{(11)}{=} \frac{1}{2}\left\|Lf^\dagger\right\|^2 + \eta\mathcal{R}(k_0)\ .
\end{aligned}$$

As $\bar{k} = k_0$ we conclude that $\bar{f}$ is a minimum-norm solution and

$$\begin{aligned}
\frac{1}{2}\left\|L\bar{f}\right\|^2 + \eta\mathcal{R}(\bar{k}) &= \lim_{m\to\infty}\left\{\frac{1}{2}\left\|Lf^m\right\|^2 + \frac{\beta_m}{\alpha_m}\mathcal{R}(k^m)\right\} \\
&= \frac{1}{2}\left\|Lf^\dagger\right\|^2 + \eta\mathcal{R}(k_0).
\end{aligned} \tag{13}$$

So far we showed the existence of a subsequence $(k^m, f^m)_m$ which converges weakly to $(k_0, \bar{f})$, where $\bar{f}$ is a minimizing solution. It remains to show that the sequence also converges in the strong topology of $\mathcal{U} \times \mathcal{V}$.

In order to show $f^m \to \bar{f}$ in $\mathcal{V}$, we prove $Lf^m \to L\bar{f}$. Since is $Lf^m \rightharpoonup L\bar{f}$ it is sufficient to show

$$\left\|Lf^m\right\| \to \left\|L\bar{f}\right\|,$$

or, as the norm is w.-l.s.c.,

$$\limsup_{m\to\infty}\left\|Lf^m\right\| \leq \left\|L\bar{f}\right\|.$$

Assume that the above inequality does not hold. Then there exists a constant τ_1 such that

$$\tau_1 := \limsup_{m\to\infty}\left\|Lf^m\right\|^2 > \left\|L\bar{f}\right\|^2$$

and there exists a subsequence of $(Lf^m)_m$ denoted by $(Lf^n)_n := (Lf^{m_n})_{m_n}$ such that

$$Lf^n \rightharpoonup L\bar{f} \quad \text{and} \quad \left\|Lf^n\right\|^2 \to \tau_1.$$

From (13) and the hypothesis stated above

$$\begin{aligned}\limsup_{n\to\infty}\frac{\beta_n}{\alpha_n}\mathcal{R}(k^n) &= \eta\mathcal{R}(k_0)+\frac{1}{2}\Big(\big\|L\bar{f}\big\|^2-\limsup_{n\to\infty}\big\|Lf^n\big\|^2\Big)\\ &< \eta\mathcal{R}(k_0),\end{aligned}$$

which is a contradiction to the w.-l.s.c. of $\mathcal{R}$. Thus

$$\limsup_{m\to\infty}\big\|Lf^m\big\| \le \big\|L\bar{f}\big\|,$$

i.e., $f^m \to \bar{f}$ in $\mathcal{V}$.

The convergence of the sequence $(k^m)_m$ in the topology of $\mathcal{U}$ follows straightforward by

$$\begin{aligned}\big\|k^m-k_0\big\| &\le \big\|k^m-k_{\epsilon_m}\big\|+\big\|k_{\epsilon_m}-k_0\big\|\\ &\le \big\|k^m-k_{\epsilon_m}\big\|+\epsilon_m \stackrel{(12)}{\longrightarrow} 0 \text{ as } m\to\infty.\end{aligned}$$

Moreover, if $f^\dagger$ is unique, the assertion about the convergence of the whole sequence $(k^j, f^j)_j$ follows from the fact that then every subsequence of the sequence converges towards the same limit $(k_0, f^\dagger)$. □

Remark 4.6. *Note that the easiest parameter choice rule fulfilling condition* (11) *is given by*

$$\beta = \eta\alpha, \qquad \eta > 0.$$

For this specific choice, we only have one regularization parameter left, and the problem (8) *reduces to*

$$\underset{(k,f)}{\text{minimize}}\;\; J_\alpha\big(k,f\big) := \frac{1}{2}T^{\delta,\varepsilon}\left(k,f\right)+\alpha\Phi(k,f)\,, \tag{14}$$

where $T^{\delta,\varepsilon}$ *is defined in* (8b) *and*

$$\Phi(k,f) := \frac{1}{2}\big\|Lf\big\|^2+\eta\mathcal{R}(k). \tag{15}$$

4.2 Convergence rates

It is well known that, under the general assumptions of the previous section, the rate of convergence of $(k^j, f^j)_j \to (k_0, f^\dagger)$ for $(\delta_j, \epsilon_j) \to 0$ will be in general arbitrarily slow. For linear and nonlinear inverse problems convergence rates were obtained if *source conditions* are satisfied [8, 9, 2, 22].

For our analysis, we will use the following source condition:

$$\mathscr{R}(B'(k_0,f^\dagger)^*) \cap \partial\Phi\big(k_0,f^\dagger\big) \neq \emptyset,$$

where $\partial\Phi$ denotes the subdifferential of the functional Φ defined in (15). This condition says there exists a subgradient $(\xi_{k_0}, \xi_{f^\dagger})$ of Φ s.t. $(\xi_{k_0}, \xi_{f^\dagger}) = B'(k_0, f^\dagger)^*\omega$, $\omega \in \mathcal{H}$.

Convergence rates are often given with respect to the *Bregman distance* generated by the regularization functional Φ. In our setting, the distance is defined by

$$D_\Phi^{(\xi_{\bar{u}},\xi_{\bar{v}})}\big((u,v),(\bar{u},\bar{v})\big) = \Phi(u,v) - \Phi(\bar{u},\bar{v}) - \big\langle(\xi_{\bar{u}},\xi_{\bar{v}})\ ,\ (u,v)-(\bar{u},\bar{v})\big\rangle \quad (16)$$

for $(\xi_{\bar{u}},\xi_{\bar{v}}) \in \partial\Phi\big(\bar{u},\bar{v}\big)$.

Lemma 4.7. *Let Φ be the functional defined in (15) with $L = I$. Then the Bregman distance is given by*

$$D_\Phi^{(\xi_{\bar{u}},\xi_{\bar{v}})}\big((u,v),(\bar{u},\bar{v})\big) = \frac{1}{2}\big\|v-\bar{v}\big\|^2 + \eta D_{\mathcal{R}}^{\zeta}\big(u,\bar{u}\big), \quad (17)$$

with $\zeta \in \partial\mathcal{R}\big(\bar{u}\big)$.

Proof. By definition of Bregman distance we have

$$\begin{aligned} D_\Phi^{(\xi_{\bar{u}},\xi_{\bar{v}})}\big((u,v),(\bar{u},\bar{v})\big) &= \left(\frac{1}{2}\|v\|^2 + \eta\mathcal{R}(u)\right) - \left(\frac{1}{2}\|\bar{v}\|^2 + \eta\mathcal{R}(\bar{u})\right) \\ &\quad - \big\langle(\xi_{\bar{u}},\xi_{\bar{v}})\ ,\ (u-\bar{u},v-\bar{v})\big\rangle \\ &= \frac{1}{2}\|v\|^2 - \frac{1}{2}\|\bar{v}\|^2 - \big\langle\xi_{\bar{v}}\ ,\ v-\bar{v}\big\rangle \\ &\quad + \eta\mathcal{R}(u) - \eta\mathcal{R}(\bar{u}) - \big\langle\xi_{\bar{u}}\ ,\ u-\bar{u}\big\rangle \\ &= \frac{1}{2}\|v-\bar{v}\|^2 + \eta D_{\mathcal{R}}^{\zeta}\big(u,\bar{u}\big) \end{aligned}$$

with $\zeta = \frac{1}{\eta}\xi_{\bar{u}}$. Note that the functional Φ is composed as a sum of a differentiable and a convex functional. Therefore, the subgradient of the first functional is an unitary set and it holds (see, e.g.,[6])

$$\begin{aligned} \partial\Phi\big(\bar{u},\bar{v}\big) &= \partial\left(\|\bar{v}\|^2 + \eta\mathcal{R}(\bar{u})\right) \\ &= \Big\{(\xi_{\bar{u}},\xi_{\bar{v}}) \in \mathcal{U}^* \times \mathcal{V}^*\ \mid\ \xi_{\bar{v}} \in \partial\|\bar{v}\|^2 \text{ and } \xi_{\bar{u}} \in \eta\partial\mathcal{R}\big(\bar{u}\big)\Big\} \end{aligned}$$

□

For the convergence rate analysis, we need the following result:

Lemma 4.8. *Let $B : \mathcal{U} \times \mathcal{V} \to \mathcal{H}$ be a bilinear operator with $\|B(k,f)\| \leq C\|k\|\|f\|$. Then its Fréchet derivative at point (k,f) is given by*

$$B'(k,f)(u,v) = B(u,f) + B(k,v),$$

$(u,v) \in \mathcal{U} \times \mathcal{V}$. *Moreover, the remainder of the Taylor expansion can be estimated by*

$$\left\|B(k+u, f+v) - B(k,f) - B'(k,f)(u,v)\right\| \leq \frac{C}{2}\left\|(u,v)\right\|^2 . \tag{18}$$

Proof. The proof is straightforward and follows from the bilinearity of the operator and its boundedness. □

The following theorem gives an error estimate within an infinite dimensional setting, similar to the results found in [17, 23]. Please note that we have not only an error estimate for the solution f, but also for the characterizing function k, i.e., we are able to derive convergence rate for the operator via (4).

Theorem 4.9 (Convergence rates). *Let $g_\delta \in \mathcal{H}$ with $\left\|g_0 - g_\delta\right\| \leq \delta$, $k_\epsilon \in \mathcal{U}$ with $\left\|k_0 - k_\epsilon\right\| \leq \epsilon$ and let $f^\dagger$ be a minimum norm solution. For the regularization parameter $0 < \alpha < \infty$, let (k^α, f^α) denote the minimizer of (14) with $L = I$. Moreover, assume that the following conditions hold:*

(i) There exists $\omega \in \mathcal{H}$ satisfying

$$(\xi_{k_0}, \xi_{f^\dagger}) = B'(k_0, f^\dagger)^* \omega,$$

with $(\xi_{k_0}, \xi_{f^\dagger}) \in \partial\Phi\big(k_0, f^\dagger\big)$.

(ii) $C\left\|\omega\right\|_{\mathcal{H}} < \min\left\{1, \frac{\gamma}{2\alpha}\right\}$, where C is the constant in (18).

Then, for the parameter choice $\alpha \sim (\delta + \epsilon)$ holds

$$\left\|B(k^\alpha, f^\alpha) - B(k_0, f^\dagger)\right\|_{\mathcal{H}} = \mathcal{O}\left(\delta + \epsilon\right)$$

and

$$D_\Phi^\xi\big((k^\alpha, f^\alpha), (k_0, f^\dagger)\big) = \mathcal{O}\left(\delta + \epsilon\right).$$

Proof. Since (k^α, f^α) is a minimizer of J_α, defined in (14), it follows

$$J_\alpha\big(k^\alpha, f^\alpha\big) \leq J_\alpha\big(k, f\big) \quad \forall (k,f) \in \mathcal{U} \times \mathcal{V}.$$

In particular,

$$\begin{aligned} J_\alpha\big(k^\alpha, f^\alpha\big) &\leq J_\alpha\big(k_0, f^\dagger\big) \\ &\leq \frac{\delta^2}{2} + \frac{\gamma\epsilon^2}{2} + \alpha\Phi(k_0, f^\dagger). \end{aligned} \tag{19}$$

Using the definition of the Bregman distance (at the subgradient $(\xi_{k_0}, \xi_{f^\dagger}) \in \partial\Phi\big(k_0, f^\dagger\big)$), we rewrite (19) as

$$\begin{aligned} &\frac{1}{2}\left\|B(k^\alpha, f^\alpha) - g_\delta\right\|^2 + \frac{\gamma}{2}\left\|k^\alpha - k_\epsilon\right\|^2 \qquad (20) \\ \leq\ & \frac{\delta^2 + \gamma\epsilon^2}{2} + \alpha\left(\Phi(k_0, f^\dagger) - \Phi(k^\alpha, f^\alpha)\right) \\ =\ & \frac{\delta^2 + \gamma\epsilon^2}{2} - \alpha\left[D_\Phi^{\xi^\dagger}\left((k^\alpha, f^\alpha), (k_0, f^\dagger)\right) + \left\langle (\xi_{k_0}, \xi_{f^\dagger})\ ,\ (k^\alpha, f^\alpha) - (k_0, f^\dagger)\right\rangle\right]. \end{aligned}$$

Using

$$\begin{aligned}\frac{1}{2}\big\|B(k^\alpha,f^\alpha)-B(k_0,f^\dagger)\big\|^2 &\le \big\|B(k^\alpha,f^\alpha)-g_\delta\big\|^2+\big\|g_\delta-g_0\big\|^2\\ &\le \big\|B(k^\alpha,f^\alpha)-g_\delta\big\|^2+\delta^2\end{aligned}$$

and

$$\begin{aligned}\frac{\gamma}{2}\big\|k^\alpha-k_0\big\|^2 &\le \gamma\big\|k^\alpha-k_\epsilon\big\|^2+\gamma\big\|k_\epsilon-k_0\big\|^2\\ &\le \gamma\big\|k^\alpha-k_\epsilon\big\|^2+\gamma\epsilon^2,\end{aligned}$$

we get

$$\begin{aligned}&\frac{1}{4}\big\|B(k^\alpha,f^\alpha)-B(k_0,f^\dagger)\big\|^2+\frac{\gamma}{4}\big\|k^\alpha-k_0\big\|^2\\ &\le \frac{1}{2}\big\|B(k^\alpha,f^\alpha)-g_\delta\big\|^2+\frac{\gamma}{2}\big\|k^\alpha-k_\epsilon\big\|^2+\left(\frac{\delta^2+\gamma\epsilon^2}{2}\right)\\ &\overset{(20)}{\le} \left(\delta^2+\gamma\epsilon^2\right)-\alpha\left[D_\Phi^{\xi^\dagger}\left((k^\alpha,f^\alpha),(k_0,f^\dagger)\right)+\left\langle(\xi_{k_0},\xi_{f^\dagger})\ ,\ (k^\alpha,f^\alpha)-(k_0,f^\dagger)\right\rangle\right].\end{aligned}$$

Denoting $r:=B(k^\alpha,f^\alpha)-B(k_0,f^\dagger)-B'(k_0,f^\dagger)((k^\alpha,f^\alpha)-(k_0,f^\dagger))$ and using the source condition *(i)*, the last term in the above inequality can be estimated as

$$\begin{aligned}&-\left\langle(\xi_{k_0},\xi_{f^\dagger})\ ,\ (k^\alpha,f^\alpha)-(k_0,f^\dagger)\right\rangle\\ &= -\left\langle B'(k_0,f^\dagger)^*\omega\ ,\ (k^\alpha,f^\alpha)-(k_0,f^\dagger)\right\rangle\\ &= \left\langle\omega\ ,\ -B'(k_0,f^\dagger)\big((k^\alpha,f^\alpha)-(k_0,f^\dagger)\big)\right\rangle\\ &= \left\langle\omega\ ,\ B(k_0,f^\dagger)-B(k^\alpha,f^\alpha)+r\right\rangle\\ &\le \|\omega\|\,\big\|B(k^\alpha,f^\alpha)-B(k_0,f^\dagger)\big\|+\|\omega\|\,\|r\|\\ &\overset{(18)}{\le} \|\omega\|\,\big\|B(k^\alpha,f^\alpha)-B(k_0,f^\dagger)\big\|+\frac{C}{2}\,\|\omega\|\,\big\|(k^\alpha,f^\alpha)-(k_0,f^\dagger)\big\|^2.\end{aligned}$$

Thus, we obtain

$$\begin{aligned}&\frac{1}{4}\big\|B(k^\alpha,f^\alpha)-B(k_0,f^\dagger)\big\|^2+\frac{\gamma}{4}\big\|k^\alpha-k_0\big\|^2+\alpha D_\Phi^{\xi^\dagger}\left((k^\alpha,f^\alpha),(k_0,f^\dagger)\right) \quad (21)\\ &\le \left(\delta^2+\gamma\epsilon^2\right)+\alpha\,\|\omega\|\,\big\|B(k^\alpha,f^\alpha)-B(k_0,f^\dagger)\big\|+\alpha\frac{C}{2}\,\|\omega\|\,\big\|(k^\alpha,f^\alpha)-(k_0,f^\dagger)\big\|^2.\end{aligned}$$

Using (17) and the definition of the norm on $\mathcal{U}\times\mathcal{V}$, (21) can be rewritten as

$$\begin{aligned}&\frac{1}{4}\big\|B(k^\alpha,f^\alpha)-B(k_0,f^\dagger)\big\|^2+\frac{\alpha}{2}\left(1-C\,\|\omega\|\right)\big\|f^\alpha-f^\dagger\big\|^2+\alpha\eta D_{\mathcal{R}}^{\zeta}(k^\alpha,k_0)\\ &\le \left(\delta^2+\gamma\epsilon^2\right)+\alpha\,\|\omega\|\,\big\|B(k^\alpha,f^\alpha)-B(k_0,f^\dagger)\big\|+\frac{1}{2}\left(\alpha C\,\|\omega\|-\frac{\gamma}{2}\right)\big\|k^\alpha-k_0\big\|^2\\ &\le \left(\delta^2+\gamma\epsilon^2\right)+\alpha\,\|\omega\|\,\big\|B(k^\alpha,f^\alpha)-B(k_0,f^\dagger)\big\|, \quad (22)\end{aligned}$$

as $\left(C\|\omega\| - \frac{\gamma}{2\alpha}\right) \le 0$ according to *(ii)*. As $(1 - C\|\omega\|)$ as well as the Bregman distance are non-negative, we derive

$$\frac{1}{4}\left\|B(k^\alpha, f^\alpha) - B(k_0, f^\dagger)\right\|^2 - \alpha\|\omega\|\left\|B(k^\alpha, f^\alpha) - B(k_0, f^\dagger)\right\| - \left(\delta^2 + \gamma\epsilon^2\right) \le 0,$$

which only holds for

$$\left\|B(k^\alpha, f^\alpha) - B(k_0, f^\dagger)\right\| \le 2\alpha\|\omega\| + 2\sqrt{\alpha^2\|\omega\|^2 + (\delta^2 + \gamma\epsilon^2)}.$$

Using the above inequality to estimate the right-hand side of (22) yields

$$\left\|f^\alpha - f^\dagger\right\|^2 \le \frac{2}{1 - C\|\omega\|}\left(\frac{\delta^2 + \gamma\epsilon^2}{\alpha} + 2\alpha\|\omega\|^2 + 2\|\omega\|\sqrt{\alpha^2\|\omega\|^2 + (\delta^2 + \gamma\epsilon^2)}\right)$$

and

$$D_{\mathcal{R}}^{\zeta}(k^\alpha, k_0) \le \frac{\delta^2 + \gamma\epsilon^2}{\eta\alpha} + \frac{2\|\omega\|}{\eta}\left(\alpha\|\omega\| + \sqrt{\alpha^2\|\omega\|^2 + (\delta^2 + \gamma\epsilon^2)}\right),$$

and for the parameter choice $\alpha \sim (\delta + \epsilon)$ follows the convergence rate $\mathcal{O}(\delta + \epsilon)$. □

Remark 4.10. *The assumptions of Theorem 4.9 include the condition*

$$C\|\omega\|_{\mathcal{H}} < \min\left\{1, \frac{\gamma}{2\alpha}\right\} .$$

Note that $\frac{\gamma}{(2\alpha)} < 1$ *for* α *small enough (i.e., for small noise level* δ *and* ϵ*), and thus* (ii) *reduces to the standard smallness assumption common for convergence rates for nonlinear ill-posed problems, see [9].*

5 A numerical example

In order to illustrate our analytical results we present first reconstructions from a convolution operator. That is, the kernel function is defined by $k_0(s,t) := k_0(s-t)$ over $\Omega = [0,1]$, see also 1 in Section 2, and we want to solve the integral equation

$$\int_\Omega k_0(s-t)f(t)dt = g_0(s)$$

from given measurements k_ϵ and g_δ satisfying (7). For our test, we defined k_0 and f as

$$k_0 = \begin{cases} 1 & x \in [0.1, 0.4] \\ 0 & \text{otherwise} \end{cases} \quad \text{and} \quad f = \begin{cases} 1 - 5|t - 0.3| & t \in [0.1, 0.5] \\ 0 & \text{otherwise} \end{cases},$$

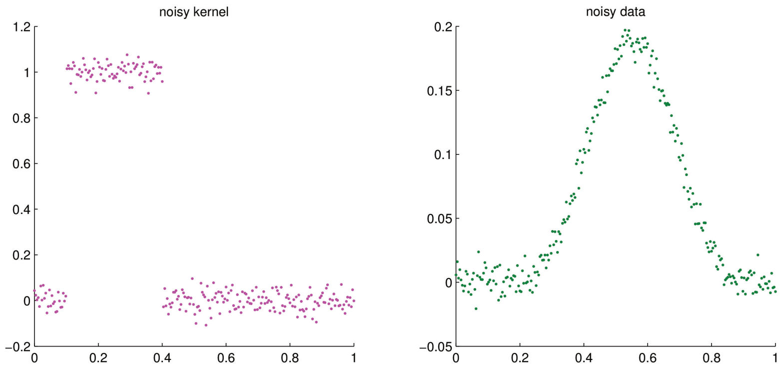

Figure 1: Simulated measurements for k_0 (left) and g_0 (right), both with 10% relative error.

respectively, the characteristic and the hat function. An example of noisy measurements k_ϵ and g_δ is displayed in Figure 1.

The functions k and f were expanded in a wavelet basis, as for example,

$$k = \sum_{l\in\mathbb{Z}} \langle k \ , \ \phi_{0,l}\rangle \phi_{0,l} + \sum_{j=0}^{\infty}\sum_{l\in\mathbb{Z}} \langle k \ , \ \psi_{j,l}\rangle \psi_{j,l} \ ,$$

where $\{\phi_\lambda\}_\lambda$ and $\{\psi_\lambda\}_\lambda$ are the pair of scaling and wavelet function associated to Haar wavelet basis. The convolution operator was implemented in terms of the wavelet basis as well. For our numerical tests, we used the Haar wavelet. The integration interval $\Omega = [0, 1]$ was discretized into $N = 2^8$ points, the maximum level considered by the Haar wavelet is $J = 6$. The functional $\mathcal{R}$ was defined as

$$\mathcal{R}(k) := \|k\|_{\ell_1} = \sum_{\lambda\in\Lambda} |\langle k \ , \ \psi_\lambda\rangle|,$$

where $\Lambda = \big\{\{l\} \cup (j, l) \ | \ j \in \mathbb{N}_0, l \le 2^j - 1\big\}$.

In order to find the optimal set of coefficients minimizing (8) we used Matlab's internal function `fminsearch`.

Figure 2 displays the numerical solutions for three different (relative) error levels: 10%, 5% and 1%. The scaling parameter was set to $\gamma = 1$ and the regularization parameters are chosen according to the noise level, i.e., $\alpha = 0.01(\delta + \varepsilon)$ and $\beta = 0.2(\delta + \varepsilon)$, ($\eta = 20$) was chosen. Our numerical results confirm our analysis. In particular it is observed that the reconstruction quality increases with decreasing noise level, see also Table 1.

Please note that the optimization with the `fminsearch` routine is by no means efficient. Currently we are working on a fast iterative optimization routine for the minimization of (8).

	$\left\|k^{\text{rec}} - k_0\right\|_1$	$\left\|f^{\text{rec}} - f^{\text{true}}\right\|_1$	$\left\|k^{\text{rec}} - k_0\right\|_2$	$\left\|f^{\text{rec}} - f^{\text{true}}\right\|_2$
10%	6.7543e-02	1.8733e-01	8.1216e-03	1.7436e-02
5%	4.0605e-02	1.7173e-01	6.9089e-03	1.5719e-02
1%	2.0139e-02	1.1345e-01	6.5219e-03	8.0168e-03

Table 1: Relative error measured by the L_1- and L_2-norm.

Acknowledgments

The authors would like to thank Dr. E. Klann, whose comments helped improving the quality of this work and Dr. A. Leitão for valuable discussion on convergence rates. The research was funded by the Austrian Science Fund (FWF): W1214-N15, project DK8.

References

[1] A. F. Boden, D. C. Redding, R. J. Hanisch, and J. Mo. Massively parallel spatially variant maximum-likelihood restoration of hubble space telescope imagery. *J. Opt. Soc. Am. A*, 13(7):1537–1545, Jul 1996.

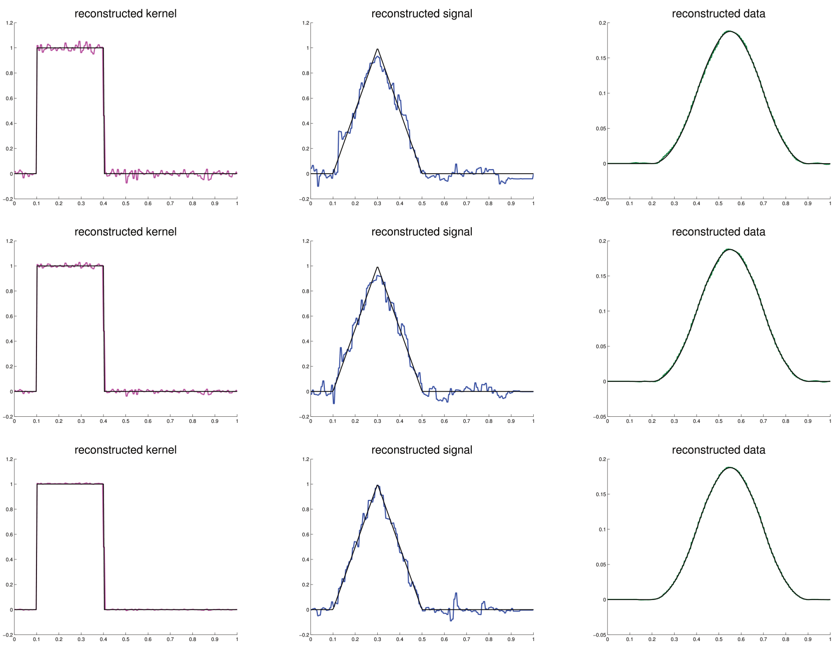

Figure 2: Reconstruction of the characterizing function k_0, the signal f (solution) and the data g_0. From top to bottom: reconstruction with 10%, 5% and 1% relative error (both for g_δ and k_ϵ). The reconstructions are colored.

[2] Martin Burger and Stanley Osher. Convergence rates of convex variational regularization. *Inverse Problems*, 20(5):1411–1421, 2004.

[3] Martin Burger and Otmar Scherzer. Regularization methods for blind deconvolution and blind source separation problems. *Math. Control Signals Systems*, 14(4):358–383, 2001.

[4] C. J. Burrows, J. A. Holtzman, S. M. Faber, P. Y. Bely, H. Hasan, C. R. Lynds, and D. Schroeder. The imaging performance of the hubble space telescope. *Astrophysical Journal*, 369:L21–L25, March 1991.

[5] Fioralba Cakoni and David Colton. *Qualitative methods in inverse scattering theory*. Interaction of Mechanics and Mathematics. Springer-Verlag, Berlin, 2006. An introduction.

[6] Frank H. Clarke. *Optimization and Nonsmooth Analysis*, volume 5 of *Classics in Applied Mathematics*. SIAM, Philadelphia, 2 edition, 1990.

[7] V. Dicken. A new approach towards simultaneous activity and attenuation reconstruction in emission tomography. *Inverse Problems*, 15(4):931–960, 1999.

[8] H W Engl, K Kunisch, and A Neubauer. Convergence rates for Tikhonov regularisation of non-linear ill-posed problems. *Inverse Problems*, 5(4):523, 1989.

[9] Heinz W. Engl, Martin Hanke, and Andreas Neubauer. *Regularization of Inverse Problems*. Kluwer Academic Publishers, Dordrecht, 1996.

[10] Gene H. Golub, Per Christian Hansen, and Dianne P. O'leary. Tikhonov regularization and total least squares. *SIAM J. Matrix Anal. Appl*, 21:185–194, 1999.

[11] Gene H. Golub and Charles F. Van Loan. An analysis of the total least squares problem. *SIAM J. Numer. Anal.*, 17(6):883–893, 1980.

[12] Jaques Hadamard. Sur les problèmes aux dérivées partielles et leur signification physique. *Princeton University Bulletin*, 13:49–52, 1902.

[13] Per Christian Hansen and Dianne P. O'leary. Regularization algorithms based on total least squares. In *in S. Van Huffel (Ed.), Recent Advances in Total Least Squares Techniques and Errors-inVariables Modeling, SIAM*, pages 127–137, 1996.

[14] L. Justen and R. Ramlau. A general framework for soft-shrinkage with applications to blind deconvolution and wavelet denoising. *Applied and Computational Harmonic Analysis*, 26(1):43 – 63, 2009.

[15] D. Kundur and D. Hatzinakos. Blind image deconvolution. *Signal Processing Magazine, IEEE*, 13(3):43 –64, may 1996.

[16] Shuai Lu and Sergei V. Pereverzev. Multi-parameter regularization and its numerical realization. *Numer. Math.*, 118(1):1–31, 2011.

[17] Shuai Lu, Sergei V. Pereverzev, and Ulrich Tautenhahn. Regularized total least squares: computational aspects and error bounds. *SIAM J. Matrix Anal. Appl.*, 31(3):918–941, 2009.

[18] Shuai Lu, Sergei V. Pereverzev, and Ulrich Tautenhahn. A model function method in regularized total least squares. *Appl. Anal.*, 89(11):1693–1703, 2010.

[19] I. Markovsky and S. Van Huffel. Overview of total least squares methods. *Signal Processing*, 87:2283–2302, 2007.

[20] F. Natterer. *The Mathematics of Computerized Tomography.* B.G. Teubner, Stuttgart, 1986.

[21] R. Ramlau. TIGRA–an iterative algorithm for regularizing nonlinear ill–posed problems. *Inverse Problems*, 19(2):433–467, 2003.

[22] Elena Resmerita. Regularization of ill-posed problems in Banach spaces: convergence rates. *Inverse Problems*, 21(4):1303–1314, 2005.

[23] U. Tautenhahn. Regularization of linear ill-posed problems with noisy right hand side and noisy operator. *Journal of Inverse and Ill-posed Problems*, 16(5):507–523, 2008.

[24] G. M. Vaĭnikko and A. Yu. Veretennikov. *Iteratsionnye protsedury v nekorrektnykh zadachakh (Extreme Methods for Solving Ill-Posed Problems).* "Nauka", Moscow, 1986.

[25] Sabine Van Huffel and Joos Vandewalle. *The total least squares problem*, volume 9 of *Frontiers in Applied Mathematics.* Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1991. Computational aspects and analysis, With a foreword by Gene H. Golub.

[26] Yu-Li You and M. Kaveh. A regularization approach to joint blur identification and image restoration. *Image Processing, IEEE Transactions on*, 5(3):416 –428, mar 1996.

Technical Reports of the Doctoral Program
"Computational Mathematics"

2012

2012-01 M.T. Khan: *Formal Semantics of* MiniMaple January 2012. Eds.: W. Schreiner, F. Winkler

2012-02 M. Kollmann, W. Zulehner: *A Robust Preconditioner for Distributed Optimal Control for Stokes Flow with Control Constraints* January 2012. Eds.: U. Langer, R. Ramlau

2012-03 W. Krendl, V. Simoncini, W. Zulehner: *Stability Estimates and Structural Spectral Properties of Saddle Point Problems* February 2012. Eds.: U. Langer, V. Pillwein

2012-04 V. Pillwein, S. Takacs: *A local Fourier convergence analysis of a multigrid method using symbolic computation* April 2012. Eds.: M. Kauers, W. Zulehner

2012-05 I. Georgieva, C. Hofreither: *Tomographic Reconstruction of Harmonic Functions* April 2012. Eds.: U. Langer, V. Pillwein

2012-06 M.T. Khan: *Formal Semantics of a Specification Language for* MiniMaple April 2012. Eds.: W. Schreiner, F. Winkler

2012-07 M. Borges-Quintana, M.A. Borges-Trenard, I. Márquez-Corbella and E. Martínez-Moro: *Computing coset leaders and leader codewords of binary codes* May 2012. Eds.: F. Winkler, V. Pillwein

2012-08 S. Takacs and W. Zulehner: *Convergence analysis of all-at-once multigrid methods for elliptic control problems under partial elliptic regularity* June 2012. Eds.: U. Langer, R. Ramlau

2012-09 C. Dönch and A. Levin: *Computation of the Strength of PDEs of Mathematical Physics and their Difference Approximations* July 2012. Eds.: F. Winkler, J. Schicho

2012-10 S. Radu and J. Sellers: *Congruences Modulo Squares of Primes for Fu's Dots Bracelet Partitions* August 2012. Eds.: M. Kauers, V. Pillwein

2012-11 I. Georgieva and C. Hofreither: *Interpolation of Harmonic Functions Based on Radon Projections* October 2012. Eds.: U. Langer, V. Pillwein

2012-12 M. Kolmbauer and U. Langer: *Efficient solvers for some classes of time-periodic eddy current optimal control problems* November 2012. Eds.: C. Pechstein, W. Zulehner

2012-13 I. Georgieva, C. Hofreither and R. Uluchev: *Least Squares Fitting of Harmonic Functions Based on Radon Projections* November 2012. Eds.: U. Langer, V. Pillwein

2012-14 G. Kiss, C. Giannelli and B. Jüttler: *Algorithms and Data Structures for Truncated Hierarchical B–splines* December 2012. Eds.: U. Langer, J. Schicho

2012-15 I. Bleyer and R. Ramlau: *A Double Regularization Approach for Inverse Problems with Noisy Data and Inexact Operator* December 2012. Eds.: E. Klann, V. Pillwein

The complete list since 2009 can be found at
`https://www.dk-compmath.jku.at/publications/`

Doctoral Program
"Computational Mathematics"

Director:

Prof. Dr. Peter Paule
Research Institute for Symbolic Computation

Deputy Director:

Prof. Dr. Bert Jüttler
Institute of Applied Geometry

Address:

Johannes Kepler University Linz
Doctoral Program "Computational Mathematics"
Altenbergerstr. 69
A-4040 Linz
Austria
Tel.: ++43 732-2468-6840

E-Mail:

office@dk-compmath.jku.at

Homepage:

http://www.dk-compmath.jku.at

Submissions to the DK-Report Series are sent to two members of the Editorial Board who communicate their decision to the Managing Editor.