REGULARIZATION METHODS FOR BLIND DECONVOLUTION AND BLIND SOURCE SEPARATION PROBLEMS

MARTIN BURGER AND OTMAR SCHERZER*

Abstract. This paper is devoted to *blind deconvolution* and *blind separation problems.* Blind deconvolution is the identification of a point spread function and an input signal from an observation of their convolution. Blind source separation is the recovery of a vector of input signals from a vector of observed signals, which are mixed by a linear (unknown) operator. We show that both problems are paradigms of nonlinear ill–posed problems. Consequently, regularization techniques have to be used for stable numerical reconstructions. In this paper we develop a rigorous convergence analysis for regularization techniques for the solution of blind deconvolution and blind separation problems. We prove convergence of the alternating minimization algorithm for the numerical solution of regularized blind deconvolution problems and present some numerical examples. Moreover, we show that many neural network approaches for blind inversion can be considered in the framework of regularization theory.

Keywords: Alternating Minimization Algorithm, Blind Deconvolution, Blind Source Separation, Ill-Posed Problems, Neural Networks, Regularization, Signal and Image Processing.

AMS Subject Classification: 35J15, 60G35, 65J20, 65K10, 68U10, 92B20.

Abbreviated Title: Regularization Methods for Blind Deconvolution.

1. Introduction. *Blind source separation* is the simultanous recovery of the matrix $K := (k_{i,j})_{1\leq i,j \leq N}$ of real valued kernel functions defined on $\mathbb{R}^d$ and the vector $\vec{x}(t) = (x_1(t), ..., x_N(t))^t$ of real-valued input signals on $\mathbb{R}^d$, from a vector of output signals $\vec{y}(s) = (y_1(s), ..., y_N(s))^t$ satisfying

$$\vec{y} = \mathcal{K}(\vec{x}, K) \; . \tag{1.1}$$

Here $\mathcal{K}(\vec{x}, K) := (\mathcal{K}_1(\vec{x}, K), ..., \mathcal{K}_N(\vec{x}, K))^t$ is a linear operator of convolution type

$$\mathcal{K}_i(\vec{x}, K) = \sum_{j=1}^{N} k_{i,j} * x_j \; , \tag{1.2}$$

and $*$ denotes convolution, i.e.,

$$k * x(s) := \int_{\mathbb{R}^d} k(s-t) x(t)\, dt \; .$$

In this paper, in order to avoid technical difficulties, we restrict our attention to the practical important situtation that the elements of the kernel matrix K are functions of convolution type. More general cases

*Institut für Industriemathematik, Johannes-Kepler-Universität Linz, A–4040 Linz, Austria. The research of the authors was supported by the Austria Fonds zur Förderung der wissenschaftlichen Forschung, F–1308 and F–1310.

can be treated using similar theoretical arguments. If $N = 1$, $d = 1$ then the simultanuous recovery of the convolution kernel k and the input signal x is usually called *blind deconvolution.*

Blind deconvolution and *blind source separation* have gained interest in many practical areas in signal- and image processing, seismology, radio-astronomy, scanning electron microscopy, medicin, and fluorescence microscopy (see e.g. [2, 3, 4, 13, 14, 16] and the references in these papers). There are actually two different philosophies for numerically solving blind deconvolution and blind source separation problems: regularization theory and neural networks. As we show in this paper neural networks can be considered as discretized regularization problems. This connection enables us to give a mathematical analysis of neural networks by utilizing the general theory of regularization methods. In classical deconvolution or restoration problems the kernel is given and only the vector $\vec{x}$ is to be determined, which implies that one has to solve a linear inverse problem. Since now the functions of the kernel matrix K and the vector $\vec{x}$ are both variables the convolution $\mathcal{K}(x, K)$ is nonlinear. Even if K is known, devonvolution can be unstable in an appropriate sense and thus blind source separation is ill–posed. This shows that blind source separation is a paradigm of a nonlinear ill–posed problem. Such problems are typically solved using regularization methods. Based on the classical theory of regularization methods for the solution of nonlinear ill–posed problems (cf. [18, 10, 9]) we give a rigorous convergence analysis of regularization methods for the solution of blind source seperation problems. We show that in the discrete setting regularization methods for blind inversion can be considered as *neural networks.*

The plan of this paper is the following: in Section 2 we introduce regularization methods for the numerical solution of blind inversion. In Section 3 we discuss various applications of blind inversion in signal-and image processing. In Section 4 we prove convergence of the alternating minimization algorithm for the solution of blind deconvolution algorithms. In Section 5 we establish the relation between neural networks and regularization methods for blind inversion and in Section 6 we present some numerical experiments.

2. Regularization Techniques for Blind Inversion. If $\vec{x}$ and K satisfy (1.1), then for any real number $c \neq 0$ $c\vec{x}$ and $\frac{1}{c}K$ satisfy (1.1), too. This obvious argument shows that blind inversion cannot be unique, and a-priori criteria have to be specified to select desired solutions. In our context we specify desired solutions which are closest to a given a-priori given pair $(\vec{x}^*, K^*)$ with respect to a Banach space norm, which will be introduced now. Let B_1 and B_2 be two Banach spaces, which are contained in $L^r(\mathbb{R}^d)$, $L^p(\mathbb{R}^d)$, respectively, with $1 \leq r, p \leq \infty$. On $(B_1^N \times B_2^{(N \times N)})$ we introduce the following norm: let $\vec{x} = (x_1, ..., x_N) \in$

B_1^N and $K = (k_{i,j})_{1\le i,j\le N} \in B_2^{(N\times N)}$. Let $\gamma > 0$, then the γ-norm on $B_1^N \times B_2^{(N\times N)}$ is defined as follows

$$\|(\vec{x}, K)\|_\gamma = \|\vec{x}\|_{B_1^N}^{\nu_{\vec{x}}} + \gamma \|K\|_{B_2^{(N\times N)}}^{\nu_K}, \tag{2.1}$$

where $\nu_{\vec{x}}$ and ν_K are two positive parameters.

The goal of this paper is to develop blind inversion algorithms which approximate a *γ–minimal norm solution* $(\vec{x}^\dagger, K^\dagger)$, that is a solution of (1.1) which is closest to $(\vec{x}^*, K^*)$ in the γ-norm.

To achieve this goal we introduce a Lagrange parameters $\lambda = 1/\alpha$ and use the minimizers of the functional

$$F_\alpha(\vec{x}, K) = \sum_{i=1}^N \|\mathcal{K}_i(x, K) - y_i\|_{L^2(\mathbf{R}^d)}^2 + \alpha \sum_{i=1}^N \|x_i - x_i^*\|_{B_1}^{\nu_x} + \alpha\gamma \sum_{i=1}^N \sum_{j=1}^N \|k_{i,j} - k_{i,j}^*\|_{B_2}^{\nu_K} \tag{2.2}$$

over $B_1^N \times B_2^{(N\times N)}$ for an approximation of the solution of (1.1). Obviously the parameter γ represents the different scaling between K and $\vec{x}$. The parameter γ should e.g. be chosen in such a way that for the desired solution $(\vec{x}^\dagger, K^\dagger)$ the relation

$$\|\vec{x}^\dagger - \vec{x}^*\|_{B_1^N}^{\nu_x} = \gamma \|K^\dagger - K^*\|_{B_2^{(N\times N)}}^{\nu_K}$$

holds. In practice, this criterion can be relaxed to

$$\|\vec{x}_\alpha^\delta - \vec{x}_*\|_{B_1^N}^{\nu_x} = \gamma \|K_\alpha^\delta - K^*\|_{B_2^{(N\times N)}}^{\nu_K},$$

where $(\vec{x}_\alpha^\delta, K_\alpha^\delta)$ is a minimizer of F_α (possibly with noisy data), which yields asymptotically the same value of γ. Another possible criterion for the choice of γ in the context of blind deconvolution is the normalization condition

$$\int_{\mathbb{R}^d} k(t)\, dt = 1, \tag{2.3}$$

which is used in applications in image processing (cf. [7]).

The existence of a minimizer of the functional F_α follows from the weak lower semi continuity of the norm in a Hilbert space and the weak lower semi continuity of the operator $\mathcal{K}$, which is proven below.

Lemma 2.1. **Continuity, weak continuity, and weak closedness of $\mathcal{K}$ in Hilbert spaces:** *Let $r = \frac{2p}{3p-2}$ and $1 \le p \le 2$.*

- *Let B_1, B_2 be two Hilbert spaces with continuous embedding in $L^p(\mathbb{R}^d)$, $L^r(\mathbb{R}^d)$, respectively. Then $\mathcal{K}$ is continuous on $B_1^N \times B_2^{(N\times N)}$.*

- *Let B_1, B_2 be two Hilbert spaces where the embedding of B_1 into $L^2(\mathbb{R}^d)$ is continuous and the embedding of B_2 into $L^1(\mathbb{R}^d)$ is absolutely continuous *. Then $\mathcal{K}$ is weakly continuous and weakly closed on $B_1^N \times B_2^{(N\times N)}$.*

Proof: It suffices to prove that each component of $\mathcal{K}$ is continuous, weakly continuous and weakly closed. Let $i \in 1, ..., N$.

First we prove that $\mathcal{K}$ is continuous. We use the fact (see e.g. [5, Thm.8.24]) that for any $k \in L^r(\mathbf{R}^d)$ and $x \in L^p(\mathbf{R}^d)$, $k * x \in L^2(\mathbf{R}^d)$ and

$$\|k * x\|_{L^2(\mathbf{R}^d)} \leq \|x\|_{L^p(\mathbf{R}^d)} \|k\|_{L^r(\mathbf{R}^d)} . \tag{2.4}$$

Therefore, for any $(\vec{x}, K)$ and $(\vec{\overline{x}}, \overline{K})$ in $B_1^N \times B_2^{(N\times N)}$, we have

$$\begin{aligned}
\|\mathcal{K}_i(x,K) - \mathcal{K}_i(\overline{x},\overline{K})\|_{L^2(\mathbb{R}^d)} &\leq \sum_{j=1}^N \|k_{i,j} * x_j - \overline{k}_{i,j} * \overline{x}_j\|_{L^2(\mathbb{R}^d)} \\
&= \sum_{j=1}^N \|k_{i,j} * x_j - k_{i,j} * \overline{x}_j + k_{i,j} * \overline{x}_j - \overline{k}_{i,j} * \overline{x}_j\|_{L^2(\mathbb{R}^d)} \\
&\leq \sum_{j=1}^N \|k_{i,j} * (x_j - \overline{x}_j)\|_{L^2(\mathbb{R}^d)} + \sum_{j=1}^N \|(k_{i,j} - \overline{k}_{i,j}) * \overline{x}_j\|_{L^2(\mathbb{R}^d)} \\
&\leq \sum_{j=1}^N \|k_{i,j}\|_{L^r(\mathbb{R}^d)} \|x_j - \overline{x}_j\|_{L^p(\mathbb{R}^d)} + \sum_{j=1}^N \|k_{i,j} - \overline{k}_{i,j}\|_{L^r(\mathbb{R}^d)} \|x_j\|_{L^p(\mathbb{R}^d)} .
\end{aligned}$$

Then, from Cauchy-Schwarz inequality it follows that

$$\begin{aligned}
\|\mathcal{K}(\vec{x},K) - \mathcal{K}(\vec{\overline{x}},\overline{K})\|^2_{L^2(\mathbb{R}^d)} \leq{}& 2\sum_{i=1}^N \sum_{j=1}^N \|k_{i,j}\|^2_{L^r(\mathbb{R}^d)} \sum_{j=1}^N \|x_j - \overline{x}_j\|^2_{L^p(\mathbb{R}^d)} \\
&+ \quad 2\sum_{i=1}^N \sum_{j=1}^N \|k_{i,j} - \overline{k}_{i,j}\|^2_{L^r(\mathbb{R}^d)} \sum_{j=1}^N \|x_j\|^2_{L^p(\mathbb{R}^d)} \\
\leq{}& \|K\|^2_{L^r(\mathbb{R}^d)^{(N\times N)}} \|\vec{x} - \vec{\overline{x}}\|^2_{L^p(\mathbb{R}^d)^N} + \|K - \overline{K}\|^2_{L^r(\mathbb{R}^d)^{(N\times N)}} \|\vec{x}\|^2_{L^p(\mathbb{R}^d)^N}
\end{aligned}$$

The continuity follows immediately from the last estimate together with the continuous embeddings of the Hilbert spaces B_1, B_2 into $L^p(\mathbb{R}^d), L^r(\mathbb{R}^d)$, respectively.

To prove weak continuity of $\mathcal{K}$ let $\vec{x}_n \rightharpoonup \vec{x}$ in B_1^N and $K_n \rightharpoonup K$ in $B_2^{(N\times N)}$. For any $\vec{z} \in L^2(\mathbf{R}^d)^N$ we have

$$\begin{aligned}
|\langle \mathcal{K}(\vec{x},K) - \mathcal{K}(\vec{x}_n,K_n), z\rangle_{L^2(\mathbb{R}^d)^N}| &\leq |\langle K * (\vec{x} - \vec{x}_n), z\rangle_{L^2(\mathbb{R}^d)^N}| + |\langle (K - K_n) * \vec{x}_n, z\rangle_{L^2(\mathbb{R}^d)^N}| \\
&\leq |\langle \vec{x} - \vec{x}_n, K^t(-.) * z\rangle_{L^2(\mathbb{R}^d)^N}| + \\
&\qquad \|\vec{x}_n\|_{L^2(\mathbb{R}^d)^N} \|K - K_n\|_{L^1(\mathbb{R}^d)^{(N\times N)}} \|z\|_{L^2(\mathbb{R}^d)^N} ;
\end{aligned}$$

*By absolutely continuous we mean that the embedding operator maps weak convergent sequence into strong convergent sequences.

where

$$K^t(-.) * z(s) := \sum_{i=1}^{N} \int_{\mathbb{R}^d} k_{i,j}(-(s-t)) z_i(t)\, dt\ .$$

Since the embedding of B_1 into $L^2(\mathbb{R}^d)$ is continuous, it is also weakly continuous (see e.g. [20]), and thus $\vec{x}_n \rightharpoonup \vec{x}$ in $L^2(\mathbb{R}^d)$. Since the embedding of B_2 into $L^1(\mathbb{R}^d)$ is absolutely continuous, we have

1. $K \in L^1(\mathbb{R}^d)^{(N\times N)}$ and thus also $K^t(-.) \in L^1(\mathbb{R}^d)^{(N\times N)}$; consequently from the general results on convolution it follows that (see e.g. [5, Thm.8.24]) $K^t(-.) * z \in L^2(\mathbb{R}^d)^N$.
2. $\|K_n - K\|_{L^1(\mathbb{R}^d)^{(N\times N)}} \to 0$.

Therefore the two terms in the last chain of inequalities tend to zero, which shows that the operator $\mathcal{K}$ is weakly continuous. Moreover, a weakly continuous operator on a Hilbert space is weakly closed. □

The existence of a minimizer follows now immediately from the weak closedness of norms in Hilbert spaces and the fact the $\mathcal{K}$ is weakly closed.

PROPOSITION 2.2. **Existence of a minimizer of F_α in Hilbert spaces:** *Let B_1, B_2 be two Hilbert spaces where the embedding of B_1 into $L^2(\mathbb{R}^d)$ is continuous and the embedding of B_2 into $L^1(\mathbb{R}^d)$ is absolutely continuous. Then F_α attains its minimum on $B_1^N \times B_2^{(N\times N)}$.*

Proposition 2.2 remains valid if we interchange the assumptions on B_1 and B_2, i.e., if the embedding of B_1 into $L^1(\mathbb{R}^d)$ is absolutely continous and the embedding of B_2 into $L^2(\mathbb{R}^d)$ is continuous, then F_α attains its minimum too.

In Banach spaces the existence of a minimizer has to be proved using a different methodology.

PROPOSITION 2.3. **Existence of a minimizer of F_α in Banach spaces:** *Let B_1 be a Banach space which is weakly closed with respect to the $L^1(\mathbb{R}^d)$-norm, i.e., if a sequence $\{\vec{x}_n\}$ of elements in B_1^N is uniformly bounded in B_1^N, then $\{\vec{x}_n\}$ has a convergent subsequence in $L^1(\mathbb{R}^d)^N$ and the limit $\vec{x}$ lies in B_1^N and satisfies*

$$\|\vec{x}\|_{B_1^N} \leq \liminf \|\vec{x}_n\|_{B_1^N}\ .$$

Moreover, let B_2 be a Hilbert space with absolutely continuous embedding in $L^1(\mathbb{R}^d)$. Then there exists a minimizer of the functional F_α on $B_1^N \times B_2^{(N\times N)}$.

Proof: Suppose that the minimum of F_α is not attained $B_1^N \times B_2^{(N\times N)}$. Then there exists a sequence $(\vec{x}_n, K_n)$ such that

$$F_\alpha(\vec{x}_n, K_n) \to \inf_{(\vec{x},K)\in B_1^N \times B_2^{N\times N}} F_\alpha(\vec{x}, K)\,, \tag{2.5}$$

and the infimum is not attained in $B_1^N \times B_2^{N\times N}$. It is an immediate consequence of (2.5) that the sequence $(\vec{x}_n, K_n)$ is uniformly bounded in $B_1^N \times B_2^{(N\times N)}$. From the assumption on B_1 it follows that $(\vec{x}_n, K_n)$ has a convergent subsequence in $(L^1(\mathbb{R}^d)^N, B_2^{(N\times N)})$, which for the sake of simplicity of notation again will be denoted by $(\vec{x}_n, K_n)$; the limit will be denoted by $(\vec{x}, K)$. Thus from the closeness assumption it follows that

$$\|\vec{x} - \vec{x}^*\|_{B_1^N} \leq \liminf \|\vec{x}_n - \vec{x}^*\|_{B_1^N} .$$

Since B_2 is a Hilbert space we have

$$\|K - K^*\|_{B_2^{(N\times N)}} \leq \liminf \|K_n - K^*\|_{B_2^{(N\times N)}} .$$

Analogously to the proof of Lemma 2.1 it follows that $\mathcal{K}(x_n, K_n) \rightharpoonup \mathcal{K}(x, K)$ in $L^2(\mathbb{R}^d)^{(N\times N)}$ and thus

$$\|\mathcal{K}(\vec{x}, K) - y\|_{L^2(\mathbb{R}^d)^{(N\times N)}} \leq \liminf \|\mathcal{K}(\vec{x}_n, K_n) - y\|_{L^2(\mathbb{R}^d)^{(N\times N)}} .$$

This shows that $(\vec{x}, K)$ is a minimum of the functional F_α, which gives a contradiction. □

Proposition 2.3 applies e.g. if B_1 is the space of functions of bounded variations (see [11]). Proposition 2.3 remains valid, if the assumptions on B_1 and B_2 are interchanged.

In the following we apply general results of regularization theory from [9, 10, 18] to prove that the minimizers of the F_α are stable in the γ-norm with respect to perturbations in the data y. The following results apply in the case that B_1 and B_2 are Hilbert spaces. In the case that B_1 or B_2 is a Banach space no general convergence and stability analysis as presented below is available.

Since by Lemma 2.1 $\mathcal{K}$ is weakly closed and continuous, the following results concerning stability and convergence of the minimizers of the functional F_α follow from general results in [10].

THEOREM 2.4. **Stability:** *Let B_1 and B_2 be two Hilbert space, where the embedding of B_1 into $L^2(\mathbb{R}^d)$ is continuous and the embedding of B_2 into $L^1(\mathbb{R}^d)$ is absolutely continuous (or where the embedding of B_2 into $L^2(\mathbb{R}^d)$ is continuous and the embedding of B_1 into $L^1(\mathbb{R}^d)$ is absolutely continuous). Moreover let $\nu_{\vec{x}} = \nu_K = 2$ (cf. 2.2). Let (δ_n) be a sequence of positive real numbers converging to 0. Let $(\vec{x}_\alpha^{\delta_n}, K_\alpha^{\delta_n})$ be a minimizer of (2.2) where the data $\vec{y}$ is replaced by some data $\vec{y}^{\delta_n}$ satisfying*

$$\|\vec{y} - \vec{y}^{\delta_n}\|_{L^2(\mathbb{R}^d)^N} \leq \delta_n . \tag{2.6}$$

Then

$$\vec{x}_\alpha^{\delta_n} \to \vec{x}_\alpha^0 \text{ in the } B_1^N - \text{norm and } K_\alpha^{\delta_n} \to K_\alpha^0 \text{ in the } B_2^{N\times N} - \text{norm} .$$

THEOREM 2.5. **Convergence:** *Let B_1 and B_2 be two Hilbert spaces, where the embedding of B_1 into $L^2(\mathbb{R}^d)$ is continuous and the embedding of B_2 into $L^1(\mathbb{R}^d)$ is absolutely continuous (or where the embedding of B_2 into $L^2(\mathbb{R}^d)$ is continuous and the embedding of B_1 into $L^1(\mathbb{R}^d)$ is absolutely continuous). Moreover let $\nu_{\vec{x}} = \nu_K = 2$. Let (δ_n), (α_n) be sequences of positive real numbers converging to 0 such that $\frac{\delta_n^2}{\alpha_n} \to 0$. Let $(\vec{x}_{\alpha_n}^{\delta_n}, K_{\alpha_n}^{\delta_n})$ be a minimizer of (2.2), where the data $\vec{y}$ is replaced by some data $\vec{y}^{\delta_n}$ satisfying (2.6) and the parameter α is replaced by α_n. Then there exists a subsequence $\tilde{\mathbb{N}}$ such that $\lim_{\tilde{n}\in\tilde{\mathbb{N}}} \vec{x}_{\alpha_{\tilde{n}}}^{\delta_{\tilde{n}}} = \vec{x}^\dagger$, where the convergence is with respect to the B_1^N-norm. Moreover, $\lim_{\tilde{n}\in\tilde{\mathbb{N}}} K_{\alpha_{\tilde{n}}}^{\delta_{\tilde{n}}} = K^\dagger$, where the convergence is with respect to the $B_2^{(N\times N)}$- norm, and $(\vec{x}^\dagger, K^\dagger)$ is a γ-minimal-norm solution.*

If, in addition, the γ-minimal-norm solution is unique, then $\lim_{\tilde{n}\in\tilde{\mathbb{N}}} \vec{x}_{\alpha_n}^{\delta_n} = \vec{x}^\dagger$ and $\lim_{\tilde{n}\in\tilde{\mathbb{N}}} K_{\alpha_n}^{\delta_n} = K^\dagger$.

It is well-known in regularization theory that the convergence of the regularized solutions $(\vec{x}_{\alpha_n}^{\delta_n}, K_{\alpha_n}^{\delta_n})$ to a γ-minimal-norm solution may be arbitraily slow (cf. Schock [17]). Therefore, investigations of conditions which guarantee convergence rates of the regularized solutions are of utmost importance. For nonlinear ill–posed problems such a result was presented in [10] first. In our context this result requires that the operator $\mathcal{K}$ is Fréchet–differentiable, with Lipschitz-continuous derivative, which is verified below.

In the following, in order to avoid notational difficulties, we restrict our attention to the case that F_α is the operator of blind deconvolution, i.e., $N = 1$ in (2.2). Consequently we set $K = k$ and $\vec{x} = x$. Analogous results for blind separation can be derived easily.

LEMMA 2.6. *Let B_1 and B_2 be two Hilbert space, where the embedding of B_1 into $L^2(\mathbb{R}^d)$ is continuous and the embedding of B_2 into $L^1(\mathbb{R}^d)$ is absolutely continuous (or where the embedding of B_2 into $L^2(\mathbb{R}^d)$ is continuous and the embedding of B_1 into $L^1(\mathbb{R}^d)$ is absolutely continuous). Let μ_1, μ_2 denote the norms of the embedding operators from B_1 into $L^2(\mathbb{R}^d)$ and from B_2 into $L^1(\mathbb{R}^d)$. Then the operator $\mathcal{K}$ is Fréchet-differentiable in $B_1 \times B_2$ with derivative*

$$\mathcal{K}'(x,k)(g,h) = g * k + x * h \,.$$

The derivative is Lipschitz-continuous and satisfies

$$\|\mathcal{K}'(x,k) - \mathcal{K}'(\overline{x},\overline{k})\| \le \sqrt{2}\mu_1\mu_2 \max\{1, \gamma^{-2}\} \|(x-\overline{x}, k-\overline{k})\|_\gamma$$

for all $x, \overline{x} \in B_1$, $k, \overline{k} \in B_2$.

Proof: Let $(g,h) \in B_1 \times B_2$. Then, from Lemma 2.1 it follows that

$$\begin{aligned}
\|\mathcal{K}(x+g,k+h) - \mathcal{K}(x,k) - \mathcal{K}'(x,k)(g,h)\|_{L^2(\mathbb{R}^d)} &= \|g*h\|_{L^2(\mathbb{R}^d)} \\
&\le \|g\|_{L^2(\mathbb{R}^d)}\|h\|_{L^1(\mathbb{R}^d)} \\
&\le \frac{\mu_1\mu_2}{2}\left(\|g\|_{B_1}^2 + \|h\|_{B_2}^2\right) \\
&\le \frac{\mu_1\mu_2}{2}\max\{1,\gamma^{-2}\}\|(g,h)\|_\gamma^2\,,
\end{aligned}$$

which implies that $\mathcal{K}'$ is the Fréchet-derivative of $\mathcal{K}$.

The Lipschitz-continuity of $\mathcal{K}'$ follows from the following estimate

$$\begin{aligned}
\|\mathcal{K}'(x,k)(g,h) - \mathcal{K}'(\overline{x},\overline{k})(g,h)\|_{L^2(\mathbb{R}^d)} &= \|(x-\overline{x})*h + g*(k-\overline{k})\|_{L^2(\mathbb{R}^d)} \\
&\le \|x-\overline{x}\|_{L^2(\mathbb{R}^d)}\|h\|_{L^1(\mathbb{R}^d)} + \|g\|_{L^2(\mathbb{R}^d)}\|k-\overline{k}\|_{L^1(\mathbb{R}^d)} \\
&\le \mu_1\mu_2\left(\|x-\overline{x}\|_{B_1} + \|k-\overline{k}\|_{B_2}\right)\|(g,h)\|_{B_1\times B_2} \\
&\le \sqrt{2}\mu_1\mu_2\|(x-\overline{x},k-\overline{k})\|_{B_1\times B_2}\|(g,h)\|_{B_1\times B_2} \\
&\le \sqrt{2}\mu_1\mu_2\max\{1,\gamma^{-2}\}\|(x-\overline{x},k-\overline{k})\|_\gamma\|(g,h)\|_\gamma.
\end{aligned}$$

$\square$

The basic ingredient in the convergence rates result in [10] is that the difference of the solution to be estimated and an a–priori given initial guess (x_*,K_*) satisfies a source wise representation of the following form:

$$(x^\dagger - x_*, k^\dagger - k_*)_\gamma = \mathcal{K}'(x^\dagger,k^\dagger)^* w_0 \tag{2.7}$$

for some $w_0 \in L^2(\mathbb{R}^d)$ with sufficiently small norm. A straightforward calculation shows that the adjoint of $\mathcal{K}'$ can be represented as

$$\mathcal{K}'(x,k)^* w = \left(E_1^*(k(-.)*w), E_2^*(x(-.)*w)\right),$$

where (E_1^*,E_2^*) denotes the adjoint of the embedding operator of $B_1\times B_2$ (where this space is equipped with the γ-norm) into $L^2(\mathbb{R}^d)$. A proof of the existence of such an operator can be found in [20]. The following result follows from direct application of the general result in [10].

Theorem 2.7. *Let B_1 and B_2 be two Hilbert space, where the embedding of B_1 into $L^2(\mathbb{R}^d)$ is continuous and the embedding of B_2 into $L^1(\mathbb{R}^d)$ is absolutely continuous (or where the embedding of B_2 into $L^2(\mathbb{R}^d)$ is continuous and the embedding of B_1 into $L^1(\mathbb{R}^d)$ is absolutely continuous). Moreover let $\nu_{\overline{x}} = \nu_K = 2$.*

Let $w_0 \in L^2(\mathbb{R}^d)$ satisfy (2.7) with $\|w_0\|_{L^2(\mathbb{R}^d)} \leq 1$, and let y^δ satisfy (2.6). Then a parameter choice $\alpha \sim \delta$ yields

$$\|x_\alpha^\delta - x^\dagger\|_{B_1}^2 + \gamma \|k_\alpha^\delta - k^\dagger\|_{B_2}^2 = O(\delta) \text{ and } \|k_\alpha^\delta * x_\alpha^\delta - y\|_{L^2(\mathbb{R}^d)} = O(\delta)\ .$$

3. Applications. In this section we apply the general results of the previous sections to some particular problems in image processing. The spaces B_1 and B_2 have to be choosen according to practical needs; in various applications, different regularization norms are required in order to apply the general convergence theory and consequently to obtain reasonable approximation properties.

3.1. Causal systems and finite-time signals. In this subsection we treat blind inversion of signals which are defined on the time interval $[0,T]$ in a causal system. For a causal system the support of the kernel function k is restricted to $\mathbb{R}^+$, which can be interpreted that the convolution takes into account data of the input signal only from the past. Blind deconvolution problems of this type occure e.g. in seismology (cf.[4]).

Appropriate function spaces for input signals are e.g. the function space $H^s([0,T])$, $s > 0$, or $BV([0,T])$. With these space we associate the function spaces $\tilde{H}^s([0,T])$ and $\tilde{BV}([0,T])$, which are functions of $H^s([0,T])$ and $BV([0,T])$, respectively, which are extended by zero outside of the intervall $[0,T]$. An appropriate space of Ansatz functions for the kernel functions is the space $L^2(\mathbb{R}^+)$. Again we consider the space $\tilde{L}^2(\mathbb{R}^+)$, which is the space of zero extensions of functions of $L^2(\mathbb{R}^+)$.

In the case $B_1 = H^s([0,T])$, $B_2 = \tilde{L}^2(\mathbb{R}^+)$ the existence of a minimizer of the functional F_α follows from Proposition 2.2. Moreover, from Theorem 2.4 stability of the minimizers of F_α with respect to data perturbations follows and from Theorem 2.5 convergence of the minimizers of F_α follows. Also the convergence rates result Theorem 2.7 is applicable. If $B_1 = BV([0,T])$ and $B_2 = \tilde{L}^2(\mathbb{R}^+)$, then Proposition 2.3 guarantees the existence of minimizer of the functional F_α.

3.2. Images with finite support. In many applications of image restoration the support Ω of the image is connected, bounded, and a–priori known. Moreover, in most applications there is also available some a–priori knowledge on the smoothness on the image to reconstruct, such as that the image to be reconstructed is in $H^s(\Omega)$ $s > 0$ or $BV(\Omega)$.

From embedding theorems (cf. [11, 15]) it follows that the embeddings of $B_1 := H^s(\Omega)$, with $s > 0$, into $L^1(\mathbb{R}^d)$ is compact – we consider again the functions of $H^s(\Omega)$ to be extended by zero outside of Ω. If we use

$B_2 = L^2(\mathbb{R}^d)$ then from Proposition 2.2 the existence of a minimizer of the functional F_α follows. Moreover, from Theorem 2.4 stability of the minimizers of F_α with respect to data perturbations follows, and from Theorem 2.5 convergence of the minimizers of F_α follows. Also the convergence rates result Theorem 2.7 is applicable. If $B_1 = BV(\Omega)$ and $B_2 = L^2(\mathbb{R}^d)$, then Proposition 2.3 guarantees the existence of minimizer of the functional F_α.

Let B_1 either $H^s(\Omega)$ with $s > 0$, or $BV(\Omega)$. Let $J_1(x) = \|x\|^2_{H^s(\Omega)}$ for $x \in B_1$ if $B_1 = H^s(\Omega)$, and $J_1(x) = \|x\|_{BV(\Omega)}$ for $x \in B_1$ if $B_1 = BV(\Omega)$.

Incorporating the a–priori information that the input image is in B_1 and the convolution kernel is in B_2, leads to the blind inversion algorithm, to minimize the functional

$$F_\alpha^{\text{fs}}(x,k) := \|k * x - y\|^2_{L^2(\mathbb{R}^d)} + \alpha\left(J_1(x) + \gamma\|k - k_*\|^2_{L^2(\mathbb{R}^d)}\right) \tag{3.1}$$

over $B_1 \times B_2$.

In the following we prove that the problem of minimizing the functional F_α^{fs} can be reduced to a nonlinear optimization problem with respect to only one variable function. The proof of the following proposition follows the presentation in [19].

PROPOSITION 3.1. *Let (x_α, k_α) be a minimizer of the functional F_α^{fs} over $B_1 \times B_2$. Then x_α minimizes the functional*

$$G_\alpha(x) = \gamma \int \frac{|\mathcal{F}x\mathcal{F}k_* - \mathcal{F}y|^2}{|\mathcal{F}x|^2 + \alpha\gamma}\, dt + J_1(x)$$

over B_1. Moreover,

$$k_\alpha = \mathcal{F}^{-1}\left(\frac{\overline{\mathcal{F}x}\mathcal{F}y + \alpha\gamma\mathcal{F}k_*}{|\mathcal{F}x|^2 + \alpha\gamma}\right) \tag{3.2}$$

Proof: Let (x_α, k_α) be a minimizer of the functional F_α^{fs} over $B_1 \times B_2$. Then (x_α, k_α) must satisy the first order optimality condition for a minimizer of the functional F_α^{fs}, which implies that for all $h \in L^2(\mathbb{R}^d)$

$$\langle k_\alpha * x_\alpha - y, h * x_\alpha\rangle_{L^2(\mathbb{R}^d)} + \alpha\gamma\,\langle k_\alpha - k_*, h\rangle_{L^2(\mathbb{R}^d)} = 0\,.$$

Combination with Plancherel's formula shows that for all $h \in L^2(\mathbb{R}^d)$

$$\langle \mathcal{F}(k_\alpha * x_\alpha - y), \mathcal{F}(h * x_\alpha)\rangle_{L^2(\mathbb{R}^d)} + \alpha\gamma\,\langle \mathcal{F}(k_\alpha - k_*), \mathcal{F}(h)\rangle_{L^2(\mathbb{R}^d)} = 0\,.$$

Thus, from [5, Thm.8.24] it follows that for all $h \in L^2(\mathbb{R}^d)$

$$\left\langle \mathcal{F}(k_\alpha)|\mathcal{F}(x_\alpha)|^2 - \mathcal{F}(y)\overline{\mathcal{F}(x_\alpha)} + \alpha\gamma\mathcal{F}(k_\alpha - k_*), \mathcal{F}(h)\right\rangle_{L^2(\mathbb{R}^d)} = 0\,. \tag{3.3}$$

Since the Fourier transform is an isomorphism on $L^2(\mathbb{R}^d)$, the set $\{\mathcal{F}(h) : h \in L^2(\mathbb{R}^d)\}$ is equal to $L^2(\mathbb{R}^d)$ and hence it follows from (3.3) that

$$\mathcal{F}(k)|\mathcal{F}(x)|^2 - \mathcal{F}(y)\overline{\mathcal{F}(x)} + \alpha\gamma\mathcal{F}(k - k_*) = 0\,,$$

which shows (3.2). Using (3.2) one shows that x_α minimizes the functional F_α^{fs} with k replaced by k_α, which is exactly the functional G_α. □

3.3. Band-limited signals and filters. In this subsection we consider blind inversion with band-limited input signals (x) and band-limited filters (k). Functions are called band–limited if the support of their spectra is contained in a compact intervall.

Let

$$F_\alpha^{\mathrm{bl}}(x, k) := \|k * x - y\|^2_{L^2(\mathbb{R})} + \alpha\,(J_1(x) + \gamma J_2(k))$$

defined on $B_1 \times B_2 := H^r(\mathbb{R}) \times H^s(\mathbb{R})$ where

$$J_1(x) = \int_{\mathbb{R}} |\mathcal{F}(x)(t)|^2(1 + |t|^{2r})dt\,, \text{ and } J_2(x) = \int_{\mathbb{R}} |\mathcal{F}(x)(t)|^2(1 + |t|^{2s})dt\,.$$

From the results in [21] it can be seen that J_1 and J_2 define equivalent norms on $H^r(\mathbb{R})$, $H^s(\mathbb{R})$, repectively.

We first consider the case of band-limited signals; the case of band–limited filters can be treated using similar arguments. In particular we assume that the support of the spectrum of the input signal to be reconstructed is in the compact intervall ω. In all what follows in this subsection, we associate to each function in $L^2(\omega)$ its extension by zero outside the compact domain ω and we will not differ between these two functions.

Then, from general properties of the Fourier–transform (see e.g. [5, Thm.8.24]) it follows that

$$F_\alpha^{\mathrm{bl}}(x, k) = G_\alpha^{\mathrm{bl}}(u, v) := \int_{\mathbb{R}} |uv - z|^2 dt + \alpha\left(\int_\omega |u|^2(1 + |t|)^{2r}dt + \gamma\int_{\mathbb{R}} |v|^2(1 + |t|^{2s})dt\right)\,,$$

where $u = \mathcal{F}(x)$, $v = \mathcal{F}(k)$, and $z = \mathcal{F}(y)$. It is easy to see that for $z \in L^2(\omega)$ any minimizer (u_α, v_α) of G_α^{bl} must satisfy $u_\alpha = 0$ on $\mathbb{R}\backslash\omega$. Hence, we can reduce the problem of minimizing G_α^{bl} over $H^r(\mathbb{R}) \times H^s(\mathbb{R})$ to the problem of minimizing G_α^{bl} over $H^r(\omega) \times H^s(\omega)$. From embedding theorems (cf. [15]) it follows that that the embedding of $B_1 := H^r(\omega)$ with $r > 0$ into $L^1(\omega)$ and the embedding of $B_2 := H^s(\omega)$ into $L^2(\omega)$ are compact. Thus from Proposition 2.2 the existence of a minimizer of the functional F_α follows. Moreover, Theorem 2.4 implies stability of the minimizers of F_α with respect to data perturbations; Theorem 2.5

implies convergence of the minimizers of F_α. Also the convergence rates result Theorem 2.7 is applicable. If $B_1 = BV(\omega)$ and $B_2 = L^2(\omega)$, then Proposition 2.3 ensures the existence of minimizer of the functional F_α^{bl}.

In the following we derive an analytical expression for the minimizers of the functional G_α^{bl}. To do so we consider the auxiliary function

$$f(p,q) := |pq - z|^2 + a|p|^2 + b|q|^2$$

defined in $\mathbb{C}^2$ with given $z \in \mathbb{C}$ and strictly positive numbers a and b. There exists an analytic expression for the minimizers of the function f as we show in the following lemma:

Lemma 3.2. *A minimizer $(\tilde{p}, \tilde{q})$ of f satisfies*

$$\tilde{p} = \frac{\overline{\tilde{q}} z}{|\tilde{q}|^2 + a}\,, \quad |\tilde{q}| = \sqrt{\max\left\{\sqrt{\frac{a}{b}}\,|z| - a, 0\right\}}\,. \tag{3.4}$$

Proof: Let $(\tilde{p}, \tilde{q})$ satisfy (3.4). Then

$$f(\tilde{p}, \tilde{q}) = \begin{cases} -ab + 2\sqrt{ab}\,|z| & \text{if } |z| \geq \sqrt{ab} \\ |z|^2 & \text{if } |z| < \sqrt{ab} \end{cases}$$

Let $|z| < \sqrt{ab}$. Then for any $(p,q) \in \mathbb{C}^2$ we have

$$\begin{aligned} |pq - z|^2 + a\,|p|^2 + b\,|q|^2 &= |pq|^2 + |z|^2 + a\,|p|^2 + b\,|q|^2 - pq\overline{z} - \overline{pq}z \\ &\geq |z|^2 + a\,|p|^2 + b\,|q|^2 - 2\,|p|\,|q|\,|z| \\ &\geq |z|^2 + 2\,|p|\,|q|\,\sqrt{ab} - 2\,|p|\,|q|\,|z| \\ &\geq |z|^2\,. \end{aligned} \tag{3.5}$$

Let $|z| \geq \sqrt{ab}$, then

$$\begin{aligned} |pq - z|^2 + a\,|p|^2 + b\,|q|^2 &= |pq|^2 + |z|^2 + a\,|p|^2 + b\,|q|^2 - pq\overline{z} - \overline{pq}z \\ &= |z|^2 - ab + (|p|^2 + b)(|q|^2 + a) - 2\,|p|\,|q|\,|z| \\ &\geq |z|^2 + (|p|^2 + b)(|q|^2 + a) - 2(|p|\,|q| + \sqrt{ab})\,|z| - ab + 2\sqrt{ab}\,|z| \qquad (3.6) \\ &\geq \left(\sqrt{(|p|^2 + b)(|q|^2 + a)} - 2(|p|\,|q| + \sqrt{ab})\right)|z| - ab + 2\sqrt{ab}\,|z| \\ &\geq -ab + 2\sqrt{ab}\,|z|\ . \end{aligned}$$

In summary we have shown that

$$f(p,q) \geq f(\tilde{p}, \tilde{q}).$$

A careful inspection of the above estimates then shows that equality in (3.5) and (3.6) only holds if (p,q) satisfies (3.4). $\square$

This lemma is the basic ingredient to calculate the given an analytical expression for the minimizer of the functional $F_\alpha^{\rm bl}$.

LEMMA 3.3. *Let $y^\delta \in L^2(\omega)$. Let $z^\delta = \mathcal{F}(y^\delta)$ denote a representative of the class of functions z^δ which is defined pointwise. Moreover, let $\alpha > 0$. Then the minimizers of F_α^{bl}, where z is replace by z^δ satisfy*

$$x_\alpha^\delta = \mathcal{F}^{-1} u_\alpha^\delta \text{ and } k_\alpha^\delta = \mathcal{F}^{-1} v_\alpha^\delta \,, \tag{3.7}$$

where u_α^δ and v_α^δ are pointwise defined by

$$\begin{aligned} u_\alpha^\delta &= sign\ (z^\delta v_*)\sqrt{\max\left\{\sqrt{\tfrac{\gamma(1+|t|^{2r})}{(1+|t|^{2s})}}\,|z^\delta| - \tfrac{\alpha\gamma}{(1+|t|^{2s})}, 0\right\}}\,, \\ v_\alpha^\delta &= sign\ (v_*)\sqrt{\max\left\{\sqrt{\tfrac{(1+|t|^{2s})}{\gamma(1+|t|^{2r})}}\,|z^\delta| - \tfrac{\alpha}{(1+|t|^{2r})}, 0\right\}}\,, \end{aligned} \tag{3.8}$$

where v_* is an arbitrary continuous function.

Proof: Since $u_\alpha^\delta = v_\alpha^\delta = 0$ in $\mathbb{R}\backslash\omega$ it follows from Cauchy-Schwarz inequality that

$$\begin{aligned} \int_\omega \frac{\left|v_\alpha^\delta\right|^2(t)}{(1+|t|^{2s})}\,dt &\le \int_\omega \sqrt{\frac{1}{\gamma(1+|t|^{2r})(1+|t|^{2s})}}|z^\delta(t)|\,dt \\ &\le \left(\int_\omega \frac{1}{\gamma(1+|t|^{2r})(1+|t|^{2s})}\,dt\right)^{\frac12}\left(\int_\omega |z^\delta(t)|^2\,dt\right)^{\frac12}\,. \end{aligned}$$

Since ω is compact, $\left(\int_\omega \frac{1}{\gamma(1+|t|^{2r})(1+|t|^{2s})}\,dt\right)^{\frac12}$ is finite, and hence

$$\int_\omega \frac{\left|v_\alpha^\delta\right|^2(t)}{(1+|t|^{2s})}\,dt \le \tilde{C}\|z^\delta\|_{L^2(\mathbb{R})}\,.$$

Similarily, we can estimate $\int_\omega \frac{\left|u_\alpha^\delta\right|^2(t)}{(1+|t|^{2s})}\,dt$. From this it follows using basic properties of the Fourier-transform that the functions x_α^δ and k_α^δ defined in (3.7) are in $H^r(\mathbb{R})$, $H^s(\mathbb{R})$, respectively. Application of Lemma 3.2 with $a = \frac{\alpha}{(1+|t|^{2r})(t)}$ and $b = \frac{\alpha\gamma}{(1+|t|^{2s})(t)}$) shows that $(u_\alpha^\delta, v_\alpha^\delta)$ is a minimizer of the functional $G_\alpha^{\rm bl}$. Moreover, it can be derived from Lemma 3.2 that for any (u,v) which differs from (3.8) on a set of positive measure the value of $G_\alpha^{\rm bl}$ is higher. This shows that $(u_\alpha^\delta, v_\alpha^\delta)$ is a minimizer of $G_\alpha^{\rm bl}$ over $L^2(\omega)\times L^2(\omega)$ if and only if it is of the form (3.8). Consequently it follows that all minimizers of the functional $F_\alpha^{\rm bl}$ over $H^r(\mathbb{R})\times H^s(\mathbb{R})$ have the form (3.7) and (3.8). □

Lemma 3.3 shows that the minimizers of the functional $F_\alpha^{\rm bl}$ are not unique. Uniqueness of the minimizer follows by requiring additional physical assumptions on k or x, like e.g., that k contains only positive frequencies.

The following theorem shows that for the particular problem of minimizing the functional $F_\alpha^{\rm bl}$ a slightly different version of the convergence theorem 2.5 can be proven.

THEOREM 3.4. *Let $\alpha,\delta\to 0$. Let $y^\delta=\mathcal{F}^{-1}z^\delta$ and let $y=\mathcal{F}^{-1}z^0$ in $L^2(\omega)$. Moroever, let v_* be a functions which is pointwise defined on $\mathbb{R}$. Let $(x_\alpha^\delta,k_\alpha^\delta)$ be as in (3.7), (3.8), respectively. Then*

$$(x_\alpha^\delta,k_\alpha^\delta)\to(\tilde{x},\tilde{k})=(\mathcal{F}^{-1}\tilde{u},\mathcal{F}^{-1}\tilde{v})$$

with respect to the $B_1\times B_2$–norm, where $\tilde{u}$ and $\tilde{v}$ satisfy

$$\begin{aligned}\tilde{u} &= \text{sign }(z^0v_*)\left(\tfrac{\gamma(1+|t|^{2r})}{(1+|t|^{2s})}\right)^{1/4}|z^\delta|^{1/2}\,,\\ \tilde{v} &= \text{sign }(v_*)\left(\tfrac{(1+|t|^{2s})}{\gamma(1+|t|^{2r})}|z^\delta|\right)^{1/4}|z^\delta|^{1/2}\,.\end{aligned}$$

Moreover $(\tilde{x},\tilde{k})=(\mathcal{F}^{-1}\tilde{u},\mathcal{F}^{-1}\tilde{v})$ is a solution of (1.1).

Proof: Let u_α^δ, v_α^δ be as in (3.7). Arguing similary as in the proof of Lemma 3.3 shows that

$$\begin{aligned}\int_\omega \tfrac{|v_\alpha^0(t)-v_\alpha^\delta(t)|^2}{(1+|t|^{2s})}\,dt &\le \int_\omega\sqrt{\tfrac{1}{\gamma(1+|t|^{2r})(1+|t|^{2s})}}|z(t)-z^\delta(t)|\,dt\\ &\le \left(\int_\omega \tfrac{1}{\gamma(1+|t|^{2r})(1+|t|^{2s})}\,dt\right)^{\frac{1}{2}}\int_\omega|z(t)-z^\delta(t)|^2\,dt\\ &\le \widetilde{C}\|z(t)-z^\delta(t)\|_{L^2(\omega)}\,.\end{aligned}$$

Moreover, we have that

$$\int_\omega\frac{\left|\tilde{v}(t)-v_\alpha^0(t)\right|^2}{(1+|t|^{2s})}\,dt\le\int_\omega\frac{\alpha}{(1+|t|^{2s})}\,ds\le\overline{C}\alpha\,.$$

Thus, it follows that

$$\begin{aligned}\|k^\dagger-k_\alpha^\delta\|^2_{H^s(\mathbb{R})}&\le 2\left(\|k^\dagger-k_\alpha\|^2_{H^s(\mathbb{R})}+\|k_\alpha-k_\alpha^\delta\|^2_{H^s(\mathbb{R})}\right)\\ &=2\left(\int_{\mathbb{R}}\frac{|v_\alpha^0-v_\alpha^\delta|^2}{(1+|t|^{2s})}\,dt+\int_{\mathbb{R}}\frac{|\tilde{v}-v_\alpha|^2}{(1+|t|^{2s})}\,dt\right)\to 0.\end{aligned}$$

Analogous arguments as used above show that $x_\alpha^\delta\to x$ in $H^r(\omega)$ and that $u_\alpha^\delta v_\alpha^\delta\to z$ in $L^2(\omega)$. From the continuity of the embedding of $H^r(\omega)$ into $L^1(\omega)$ it follows from general results on the convolution that $k_\alpha^\delta * x_\alpha^\delta\to z$.

Finally we note that it is easy to see that $(\tilde{x},\tilde{k})$ is a solution of (1.1). □

We note that in contrast to Theorem 2.5 we have not required an assumption like $\frac{\delta^2}{\alpha}\to 0$ as $\alpha,\delta\to 0$. However, we were not able to prove convergence to a solution of minimal norm.

4. The alternating minimization algorithm for blind deconvolution. For blind deconvolution the functional F_α is convex in both components x and k. In fact keeping either one of the components x or k fixed the resulting minimization problem can be reduced to the solution of a linear equation. This makes an alternating minimization algorithm a reasonable candidate for numerically solving blind inversion problems (cf. [3, 7, 13]). To be precise: let $k^0 \in B_2$. Then the alternating minimization algorithm consists in alternate calculation of a minimizer x_{n+1} of $F_\alpha(x, k_n)$ on B_1 (for given k_n) and of a minimizer k_{n+1} of $F_\alpha(x_{n+1}, k)$.

In the following we establish well-definedness and a weak convergence result of the alternating minimization algorithm.

LEMMA 4.1. *Let B_1 be a Hilbert-space which can be continuously embedded in $L^2(\mathbb{R}^d)$ and let B_2 be a Banach-space which is weakly closed with respect to the $L^1(\mathbb{R}^d)$-norm (cf. Proposition 2.3) (or let B_2 be a Hilbert space which can be compactly embedded in $L^1(\mathbb{R}^d)$).*

Then every step of the alternating minimization algorithm is well–defined, and there exist a subsequence $(n^l) \subset \mathbb{N}$ and functions $\overline{x}, \tilde{x} \in L^2(\mathbb{R}^d)$, $\overline{k} \in L^1(\mathbb{R}^d)$, such that

$$x_{n^l} \rightharpoonup \overline{x} \text{ in } B_1\,, \quad x_{n^l+1} \rightharpoonup \tilde{x} \text{ in } B_1, \text{ and } k_{n^l} \to \overline{k} \text{ in } L^1(\mathbb{R}^d)\,.$$

Proof: From Lemma 2.1 the well–definedness of the alternating minimization algorithm follows if B_2 is a Hilbert-space. Arguing as in the proof of Proposition 2.3 shows the well–definedness of the alternating minimization algorithm if B_2 is a Banach-space which is weakly closed with respect to the $L^1(\mathbb{R}^d)$-norm. From the basic estimates

$$\begin{aligned}
\alpha\|x_n - x_*\|_{B_1}^{\nu_x} + \alpha\gamma\|k_n - k_*\|_{B_2}^{\nu_K} &\le F_\alpha(x_n, k_n) \\
&= \min_{k\in B_2} F_\alpha(x_n, k) \\
&\le F_\alpha(x_n, 0) \\
&= \|y\|^2_{L^2(\mathbb{R}^d)} + \alpha\|x_n - x_*\|_{B_1}^{\nu_x} + \alpha\gamma\|k_*\|_{B_2}^{\nu_K}
\end{aligned}$$

and

$$\begin{aligned}
\alpha\|x_{n+1} - x_*\|_{B_1}^{\nu_x} + \alpha\gamma\|k_n - k_*\|_{B_2}^{\nu_K} &\le F_\alpha(x_{n+1}, k_n) \\
&= \min_{x\in B_1} F_\alpha(x, k_n) \\
&\le F_\alpha(0, k_n) \\
&\le \|y\|^2_{L^2(\mathbb{R}^d)} + \alpha\|x_*\|_{B_1}^{\nu_x} + \alpha\gamma\|k_n - k_*\|_{B_2}^{\nu_K}
\end{aligned}$$

it follows that

$$\|x_n - x_*\|_{B_1} \le \left(\frac{1}{\alpha}\|y\|^2_{L^2(\mathbb{R}^d)} + \|x_*\|_{B_1}^{\nu_x}\right)^{1/\nu_x} \text{ and } \|k_n - k_*\|_{B_2} \le \left(\frac{1}{\alpha\gamma}\|y\|^2_{L^2(\mathbb{R}^d)} + \|k_*\|_{B_2}^{\nu_K}\right)^{1/\nu_K}\,.$$

Hence, there exists $I \subset \mathbb{N}$ such that $x_i \rightharpoonup \overline{x}$ in B_1. Since the embedding of B_2 into $L^1(\mathbb{R}^d)$ is compact, we can choose I such that also $k_i \to \overline{k}$ in $L^1(\mathbb{R}^d)$. With the same argument it follows that there exists a subsequence $\{x_{n^l}\} \subset \{x_i\}$ and $\tilde{x} \in B^1$ such that $x_{n^l+1} \rightharpoonup \tilde{x}$ in B_1. □

In the following we establish that a subsequence of the iterates of the alternating minimization algorithm is strongly convergent.

LEMMA 4.2. *Let B_1 be a Hilbert-space which can be continuously embedded in $L^2(\mathbb{R}^d)$ and let B_2 be a Banach-space which is weakly closed with respect to the $L^1(\mathbb{R}^d)$-norm (cf. Proposition 2.3) (or let B_2 be a Hilbert space, which can be compactly embedded in $L^1(\mathbb{R}^d)$). Let (n^l) be a subsequence of $\mathbb{N}$ such that $x_n^l \rightharpoonup \overline{x}$ in the B_1–norm. Moreover, let $k_n^l \to \overline{k}$ in the $L^1(\mathbb{R}^d)$–norm. Then $x_{n^l+1} \to \tilde{x}$ in B_1.*

Proof: Since $\nu_x = 2$ we have that for fixed $k_{n^l} \in B_2$, the function $F_\alpha(., k_{n^l})$ is quadratic.

Thus a function x_{n^l+1} minimizes the functional $F_\alpha(., k_{n^l})$ if and only if it satisfies the first order optimality condition

$$\left\langle k_{n^l} * x_{n^l+1}, k_{n^l} * x \right\rangle_{L^2(\mathbb{R}^d)} + \alpha \left\langle x_{n^l+1} - x_*, x \right\rangle_{B_1} = \left\langle y, k_n^l * x \right\rangle_{L^2(\mathbb{R}^d)} . \tag{4.1}$$

Since by assumption $x_{n^l+1} \rightharpoonup \tilde{x}$ with respect to the B_1–norm we get $\left\langle x_{n^l+1}, x \right\rangle_{B_1} \to \left\langle \tilde{x}, x \right\rangle_{B_1}$. From elementary properties of the Fourier-transform (as used already before) we get

$$\left| \left\langle y, (k_n^l - \overline{k}) * x \right\rangle_{L^2(\mathbb{R}^d)} \right| \le \|y\|_{L^2(\mathbb{R}^d)} \|x\|_{L^2(\mathbb{R}^d)} \|k_{n^l} - \overline{k}\|_{L^1(\mathbb{R}^d)} .$$

The last term tends to zero for $n^l \to \infty$ since we assumed that $k_n \to \overline{k}$ with respect to the in $L^1(\mathbb{R}^d)$–norm. Moreover, we have

$$\begin{aligned}
&| \left\langle k_n^l * x_{n^l+1}, k_n^l * x \right\rangle - \left\langle \overline{k} * \tilde{x}, \overline{k} * x \right\rangle_{L^2(\mathbb{R}^d)} | \\
&\le | \left\langle k_n^l * x_{n^l+1}, (k_n^l - \overline{k}) * x \right\rangle_{L^2(\mathbb{R}^d)} | + | \left\langle (k_n^l - \overline{k}) * x_{n^l+1}, \overline{k} * x \right\rangle_{L^2(\mathbb{R}^d)} | + | \left\langle \overline{k} * (x_{n^l+1} - \tilde{x}), \overline{k} * x \right\rangle_{L^2(\mathbb{R}^d)} | \\
&\le \|k_{n^l} - \overline{k}\|_{L^1(\mathbb{R}^d)} \|x_{n^l+1}\|_{L^2(\mathbb{R}^d)} \|x\|_{L^2(\mathbb{R}^d)} \left(\|k_{n^l}\|_{L^1(\mathbb{R}^d)} + \|\overline{k}\|_{L^1(\mathbb{R}^d)} \right) + \quad | \left\langle \overline{k} * (x_{n^l+1} - \tilde{x}), \overline{k} * x \right\rangle_{L^2(\mathbb{R}^d)} | .
\end{aligned}$$

Since $k_{n^l} \to \overline{k}$ in the L^1–norm and since $\|x_{n^l+1}\|_{L^2(\mathbb{R}^d)}$ and $\|k_{n^l+1}\|_{L^1(\mathbb{R}^d)}$ are uniformly bounded, the first term on the right hand side in the last chain of inequalities converges to zero. The second term on the right hand side in the last chain of inequalities converges to zero since $x_{n^l+1} \rightharpoonup \tilde{x}$ in the B_2–norm. Since the embedding of B_2 into $L^2(\mathbb{R}^d)$ is continuous, it is also weakly continuous, and hence $x_{n^l+1} \rightharpoonup \tilde{x}$ in the $L^2(\mathbb{R}^d)$–norm. Therefore, it follows from (4.1) that

$$\left\langle \overline{k} * \tilde{x}, \overline{k} * x \right\rangle_{L^2(\mathbb{R}^d)} + \alpha \left\langle \tilde{x} - x_*, x \right\rangle_{B_1} = \left\langle y, \overline{k} * x \right\rangle_{L^2(\mathbb{R}^d)} . \tag{4.2}$$

Using $x = x_{n^l+1} - \overline{x}$ in (4.1) and (4.2) we get

$$\begin{aligned}
&\|\overline{k} * (\tilde{x} - x_{n^l+1})\|^2_{L^2(\mathbb{R}^d)} + \alpha \|\tilde{x} - x_{n^l+1}\|^2_{B_1} \\
&= \quad \left\langle \overline{k} * (\tilde{x} - x_{n^l+1}), \overline{k} * (\tilde{x} - x_{n^l+1}) \right\rangle_{L^2(\mathbb{R}^d)} + \alpha \left\langle \tilde{x} - x_{n^l+1}, \tilde{x} - x_{n^l+1} \right\rangle_{B_1} \\
&= \quad \left\langle (k_n^l - \overline{k}) * x_{n^l+1}, k_n^l * (\tilde{x} - x_{n^l+1}) \right\rangle_{L^2(\mathbb{R}^d)} + \\
&\quad \left\langle \overline{k} * x_{n^l+1}, (k_{n^l} - \overline{k}) * (\tilde{x} - x_{n^l+1}) \right\rangle_{L^2(\mathbb{R}^d)} + \\
&\quad \left\langle y, (k_n^l - \overline{k}) * (\tilde{x} - x_{n^l+1}) \right\rangle_{L^2(\mathbb{R}^d)} .
\end{aligned}$$

Let μ_1 denote the norm of the embedding operator of B_1 into $L^2(\mathbb{R}^d)$, then from basic properties of the convolution (as used already before) it follows that

$$\|\tilde{x} - x_{n^l+1}\|_{B_1} \leq \mu_1 \frac{\|k_n^l - \overline{k}\|_{L^1(\mathbb{R}^d)}}{\alpha} \left(\left[\|\overline{k}\|_{L^1(\mathbb{R}^d)} + \|k_n^l\|_{L^1(\mathbb{R}^d)} \right] \|x_{n^l+1}\|_{L^2(\mathbb{R}^d)} + \|y\|_{L^2(\mathbb{R}^d)} \right) .$$

From this and the assumption that $\|k_n^l - \overline{k}\|_{L^1(\mathbb{R}^d)} \to 0$ it follows that $\|\tilde{x} - x_{n^l+1}\|_{B_1} \to 0$. □

We will use Lemma 4.2 to show that the alternating minimization algorithm is convergent (and not only a subsequence as in Lemma 4.2) if B_2 is a Hilbert–space.

THEOREM 4.3. *Let B_1 be a Hilbert-space which can be compactly embedded in $L^2(\mathbb{R}^d)$ and let B_2 be a Hilbert space, which can be compactly embedded in $L^1(\mathbb{R}^d)$. Moreover, let in the functional F_α $\nu_x = \nu_K = 2$. Then*

$$x_n \to \overline{x} \textit{ in the } B_1 - \textit{ norm and } k_n \to \overline{k} \textit{ in the } L^1(\mathbb{R}^d) - \textit{ norm },$$

where $(\overline{x}, \overline{k}) \in B_1 \times B_2$ satisfies for all $x \in B_1$ and $k \in B_2$

$$F_\alpha(\overline{x}, \overline{k}) \leq F_\alpha(\overline{x}, k) \textit{ and } F_\alpha(\overline{x}, \overline{k}) \leq F_\alpha(x, \overline{k}) .$$

Proof: From Lemma 4.1 and the assumption that B_1 can be compactly embedded into $L^2(\mathbb{R}^d)$ it follows that there exists a subsequence (n^l) of $\mathbb{N}$ satisfying $x_{n^l} \to \overline{x}$ in B_1, $x_{n^l+1} \to \tilde{x}$ in B_1, and $k_{n^l} \to \overline{k}$ in $L^1(\mathbb{R}^d)$. Then, from Lemma 4.2 it follows that $x_{n^l+1} \to \tilde{x}$ in the B_1–norm (actually there exists only a subsequence - in order to avoid notational difficulties we give the subsequence the same name). From the definition of the alternating minimization algorithm it follows that

$$F_\alpha(x_{n^{l+1}}, k_{n^{l+1}}) \leq \cdots \leq F_\alpha(x_{n^l+1}, k_{n^l}) \leq F_\alpha(x_{n^l}, k_{n^l}) .$$

Using this and the continuity of the functional F_α on $B_1 \times B_2$ (cf. Lemma 2.1) it follows that

$$F_\alpha(\tilde{x}, \overline{k}) = F_\alpha(\overline{x}, \overline{k}) .$$

Since the minimizer of F_α is unique, for fixed $\overline{k}$ it follows that $\tilde{x} = \overline{x}$. Now the assertion follows from a standard subsequence of subsequences argument. $\square$

Theorem 4.3 shows that the alternating minimization algorithm is convergent although the minimizers of the functional F_α are not unique. In order to explain the behaviour of the alternating minimization algorithm we shift $x(t)$ to $x(t-\tau)$ and $k(t)$ to $k(t+\tau)$. The convolution is invariant with respect to this shift, i.e. $x * k = x_0 * k_0$. Moreover, all the regularization norms, like e.g. the $H^s(\mathbb{R}^d)$ norm, are also shift invariant. It is now a simple observation that if x minimizes $F_\alpha(.,k)$ (for fixed k), then $x(.-\tau)$ is the minimizer of $F_\alpha(.,k(.+\tau))$. This implies that if start the alternating minimization algorithm with $k_0(.+\tau)$, then the iterates are just the translations $(x_n(.-\tau), k_n(.+\tau))$ of the iterates from the alternating minimization algorithm with initial value k_0.

5. Numerical Implementation and Neural Networks. In practical applications only a discrete data set $\{y_j = y(t_j)\}_{j=1,\dots,N}$ of the continuous data y is available. In this case it is natural to consider instead of minimizing the functional F_α the modified functional F_α^{d} involving only the available discrete data

$$F_\alpha^{\mathrm{d}}(x,k) := \frac{1}{N}\sum_{j=1}^{N}(k * x(t_j) - y(t_j))^2 + \alpha\|x - x_*\|_{B_1}^{\nu_x} + \alpha\gamma\|k - k_*\|_{B_2}^{\nu_K} \ .$$

over $B_1 \times B_2$.

For the sake of simplicity of notation we set in this section $x_* = k_* = 0$, $\nu_x = \nu_K = 2$ and use $B_1 = H^r(\mathbb{R}^d)$ and $B_2 = H^s(\mathbb{R}^d)$, then with appropriate differential operators L_1 and L_2 the inner products of these spaces can be rewritten as

$$\langle x, \overline{x}\rangle_{B_1} = \langle L_1 x, L_1\overline{x}\rangle_{L^2(\mathbb{R}^d)} \ , \quad \langle k, \overline{k}\rangle_{B_2} = \langle L_2 k, L_2\overline{k}\rangle_{L^2(\mathbb{R}^d)} \ .$$

For these space the alternating minimization algorithm can be realized in two steps:

First we consider the minimization of F_α^{d} in x for fixed k. For fixed k the minimizer of F_α^{d} must satisfy the first order optimality condition

$$\frac{1}{N}\sum_{j=1}^{N}(k * x(t_j) - y_j)k * h(t_j) + \alpha\int_{\mathbb{R}} L_1 x L_1 h \ .$$

Therefore with the notation

$$c_j := \frac{k * x(t_j) - y_j}{N\alpha}$$

it follows that

$$L_1^* L_1 x(t) = -\sum_{j=1}^{N} c_j \ k(t_j - t) \, ,$$

where L_1^* denotes the L^2-adjoint operator of L_1. Therefore

$$x(t) = \sum_{j=1}^{N} c_j \; X(t_j - t)\,, \tag{5.1}$$

with $X = -(L_1^* L_1)^{-1} k$ and the coefficients c_j satisfy

$$\begin{aligned} N\alpha c_j &= x * k(t_j) - y_j \\ &= \sum_{m=1}^{N} c_m (X(t_m - .) * k)(t_j) - y_j \\ &= \sum_{m=1}^{N} c_m \int X(t_m - t) k(t_j - t)\; dt - y_j \\ &= \sum_{m=1}^{N} c_m \int X(t_m - t) L_1^* L_1 X(t_j - t)\; dt - y_j \\ &= -\sum_{m=1}^{N} c_m \int L_1 X(t_m - t) L_1 X(t_j - t)\; dt - y_j\,. \end{aligned}$$

Therefore the calculation of the vector $c = (c_1, ..., c_N)$ reduces to the solution of the matrix equation

$$(A + N\alpha I)c = \mathbf{y}\,, \tag{5.2}$$

where A is the positively definite matrix

$$A = (A_{jm}) = \int L_1 X(t_m - t) L_1 X(t_j - t)\; dt\,.$$

Secondly we consider the minimization of F_α^{d} in K for fixed x. Analogously as above we find from the optimality condition for x that

$$k(t) = \sum_{j=1}^{N} d_j \; K(t_j - t)\,, \tag{5.3}$$

where $K = (L_2^* L_2)^{-1} x$ and the vector d is the solution of the matrix equation

$$(B + N\alpha\gamma I)d = \mathbf{y}, \tag{5.4}$$

with the positively definite matrix

$$B = (B_{jm}) = \int L_2 K(t_m - t) L_2 K(t_j - t)\; dt\,.$$

Thus one step of the alternating minimization algorithm (which is repeated iteratively) can be summarized as follows:

- Construct the basis function X_{n+1} by solving $L_1^* L_1 X_{n+1} = -k_n$.
- Allocate the matrix A in (5.2) with X replaced by X_{n+1}.
- Compute the coefficients c_j by solving (5.2).

- Compute $x_{n+1} = \sum_{j=1}^{N} c_j \; X_{n+1}(t_j - t)$.
- Construct the basis function K_{n+1} by solving $L_2^* L_2 K_{n+1} = -x_{n+1}$.
- Allocate the matrix B with K replaced by K_{n+1}.
- Compute the coefficients d_j by solving (5.4).
- Compute $k_{n+1} = \sum_{j=1}^{N} d_j \; K_{n+1}(t_j - t)$.

Given values of an input signal x at some samples $t_0 < t_1 < ... < t_N$ neural networks are used to predict function values of x at time $t_{\text{pred}} > t_N$ and $t_{\text{pred}} < t_0$. Neural networks are based on the assumption that the underlying physical model is convolution, i.e., $y = k * x$ where y is given data and the kernel function k is unknown too.

From (5.1) and (5.3) it follows that

$$(L_1^* L_1)(L_2^* L_2)x(t) = -\sum_{j=0}^{N}\sum_{i=0}^{N} c_j x(t_i - t_j + t) d_i \; . \tag{5.5}$$

Once the vectors c and d are calculated, this model can be used for prediction, as we illustrate with the following example: Let $L_1 = L_2$ be the first derivative, then (5.5) becomes

$$x^{IV}(t) = -\sum_{j=0}^{N}\sum_{i=0}^{N} c_j x(t_i - t_j + t) d_i \; . \tag{5.6}$$

Assume for simplicity of notation that $t_i = i$. Then, writing (5.6) at discrete sampling points gives

$$\begin{pmatrix} x^{IV}(0) \\ .... \\ x^{IV}(N) \end{pmatrix} = (c_0, ..., c_N) \begin{pmatrix} x(0) & x(-1) & ... & x(-N) \\ x(1) & x(0) & ... & x(-N+1) \\ ... & ... & ... & ... \\ x(N) & x(N-1) & ... & x(0) \end{pmatrix} \begin{pmatrix} d(0) \\ ... \\ ... \\ d(N) \end{pmatrix}$$

The last line of the matrix equation allows to calculate the fourth derivative of x at t_N. In neural networks this information is used for extrapolating the function value at t_{N+1} via Taylor series expansion

$$x(t_{N+1}) \sim x(t_N) + x^I(t_N) + \frac{1}{2}x^{II}(t_N) + \frac{1}{6}x^{III}(t_N) + \frac{1}{24}x^{IV}(t_N) \; .$$

Note that the third, second, and first derivatives at t_N can be determined inductively via

$$x^{III}(t_i) - x^{III}(t_0) = \int_{t_0}^{t_i} x^{IV}(s)\, ds \; \sim \sum_{j=1}^{i} x^{IV}(t_j)(t_j - t_{j-1}) \, ,$$

and similar formulas for the first and second derivatives. If instead of regularizing with both the H^1-norm we would have regularized the kernel k with the L^2-norm and the input signal with the H^1-norm, then we can use the first and second derivative at t_N for extrapolation. This shows that the regularization norm determine derivatives at the boundary of the available measured data. Each regularization norm gives raise to a different neural network. The process of extrapolation of future data $t > t_{N+1}$ can be continued iteratively, by just using the data of x at $t_2, ..., t_{N+1}$ instead of $t_1, ..., t_N$ with the same vectors c and d. In practice, control measurements of $y(t)$ are performed and the predictions will be compared with the control data. Once the prediction becomes inaccurate, more data y is used to determine updated vectors c_{new} and d_{new} and the whole process is repeated.

6. Numerical Results. In the following we present some results of blind deconvolved signals in $\mathbb{R}^1$, which illustrate the results obtained in the preceding sections. Many other numerical results for blind deconvolution problems in $\mathbb{R}^1$ and $\mathbb{R}^2$ can be found in literature (cf. e.g [3, 7, 13]). In our numerical test examples presented below the kernel function to be recovered is

$$k(t) = ce^{-\frac{t^2}{2}} , \tag{6.1}$$

except for the ones in Figure 6.4, where the point spread function k to be recovered is

$$k = \frac{\sin x}{x}, \tag{6.2}$$

whose fourier transform is a characteristic function in the scale space. The images to be reconstructed are a piecewise cubic spline in Figures 6.1, 6.4, 6.5 and 6.6, a piecewise quadratic peak in Figure 6.2 and a Gaussian in Figure 6.3. For stabilization we used in F_α the terms $\|x\|^2_{L^2(\mathbb{R})}$ and $\|k\|^2_{L^2(\mathbb{R})}$. The parameter γ is chosen such that (2.3) holds, which also agrees with the special choice of the PSF functions. Moreover, we use the a-priori knowledge that the PSF is symmetric and that only physically meanigful frequencies occur, i.e., that $\mathcal{F}(k) \geq 0$. In the context of Theorem 3.4, i.e., for band-limited signals or filters, this a-priori assumption implies the uniqueness of a minimizer of F_α.

The data are generated synthetically by specifying the functions x and k, which makes it possible to compare the reconstructions obtained by the blind inversion algorithm with the exact solutions. In all figures, the exact image resp. PSF is drawn as a solid line, the computed approsimation is drawn as a solid line with additional dots.

All results show that the supports of both the input signal and the kernel function can be recovered reasonably accurate with the blind inversion algorithm. However, the L^2-stabilization term leads to significant smoothing, the 'mass' is distributed around the support but the maximum is lower, which can be e.g. realized in the case of the peak-functions in Figure 6.1 and 6.2. In Figure 6.3, the image x is chosen to be a

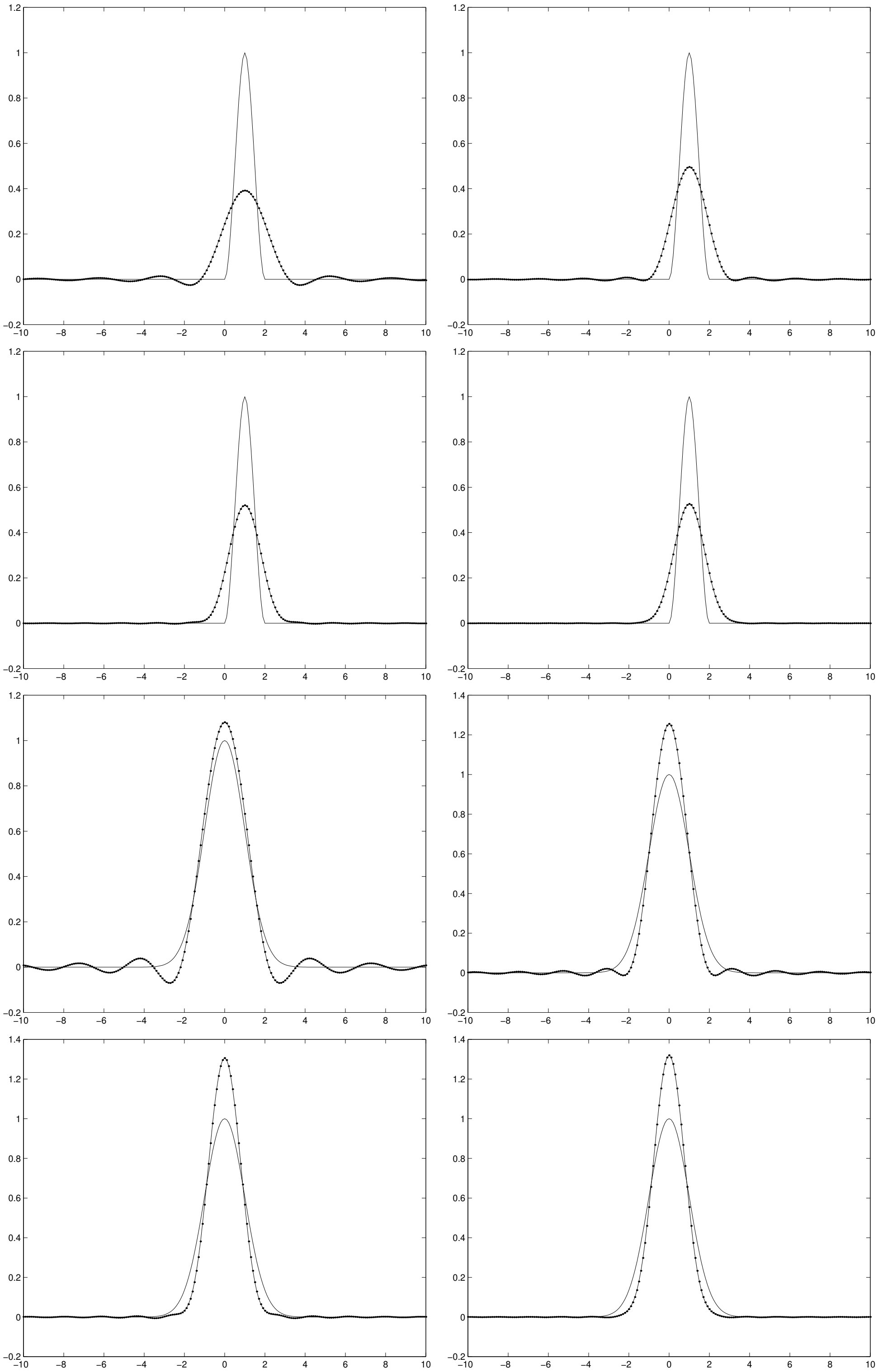

FIG. 6.1. *Reconstruction of a piecewise cubic spline image (first four pictures) and a gaussian kernel (pictures below) for the parameter values* $\alpha = 0.1, 0.01, 0.001, 0.0001$.

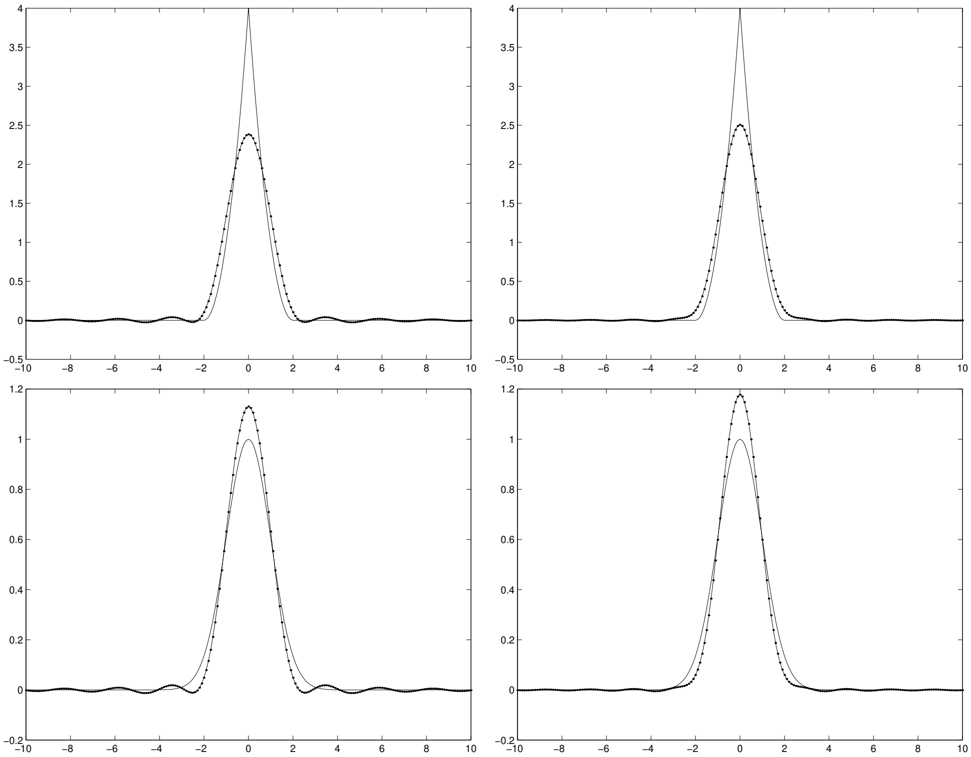

FIG. 6.2. *Reconstruction of a peak (above) and a gaussian PSF (below) for $\alpha = 0.01$ (left) and $\alpha = 0.001$.*

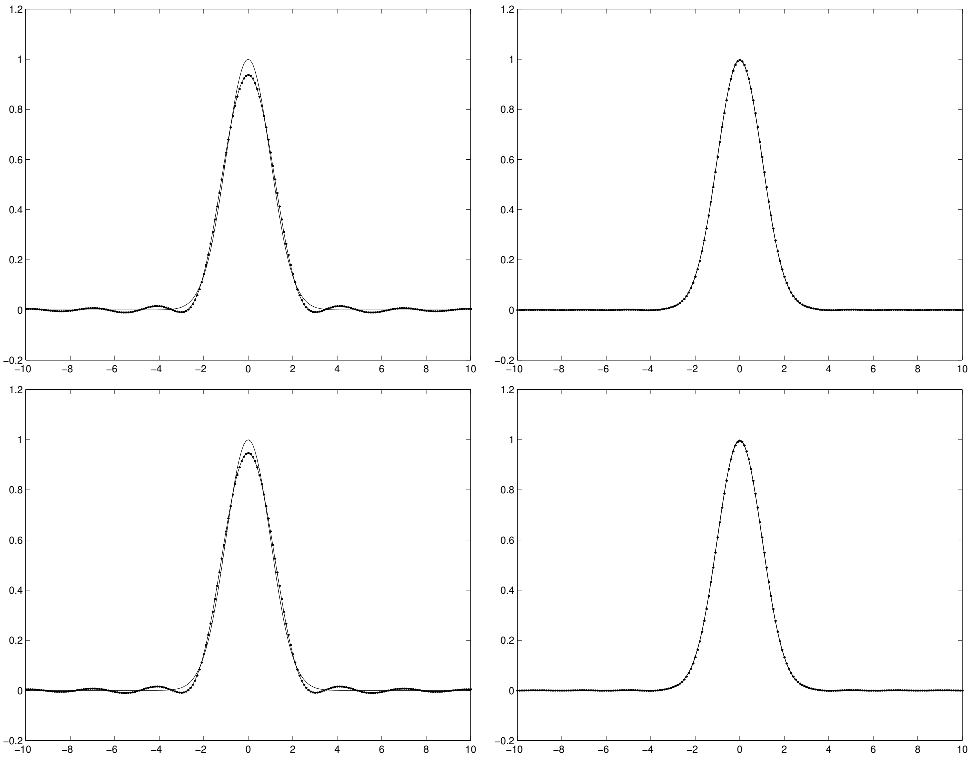

FIG. 6.3. *Reconstruction of a gaussian image (above) and a gaussian PSF (below) for $\alpha = 0.01$ (left) and $\alpha = 0.0001$.*

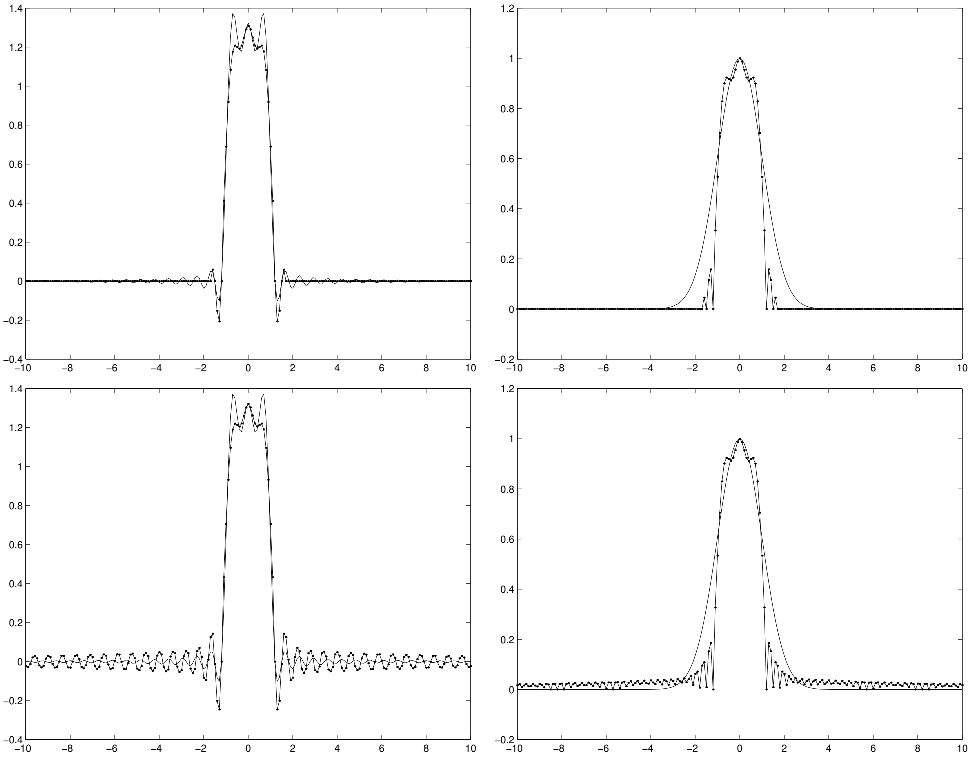

FIG. 6.4. *Reconstruction in the scale space of a smooth signal and a filter function for $\alpha = 0.01$ and $\alpha = 0.0001$.*

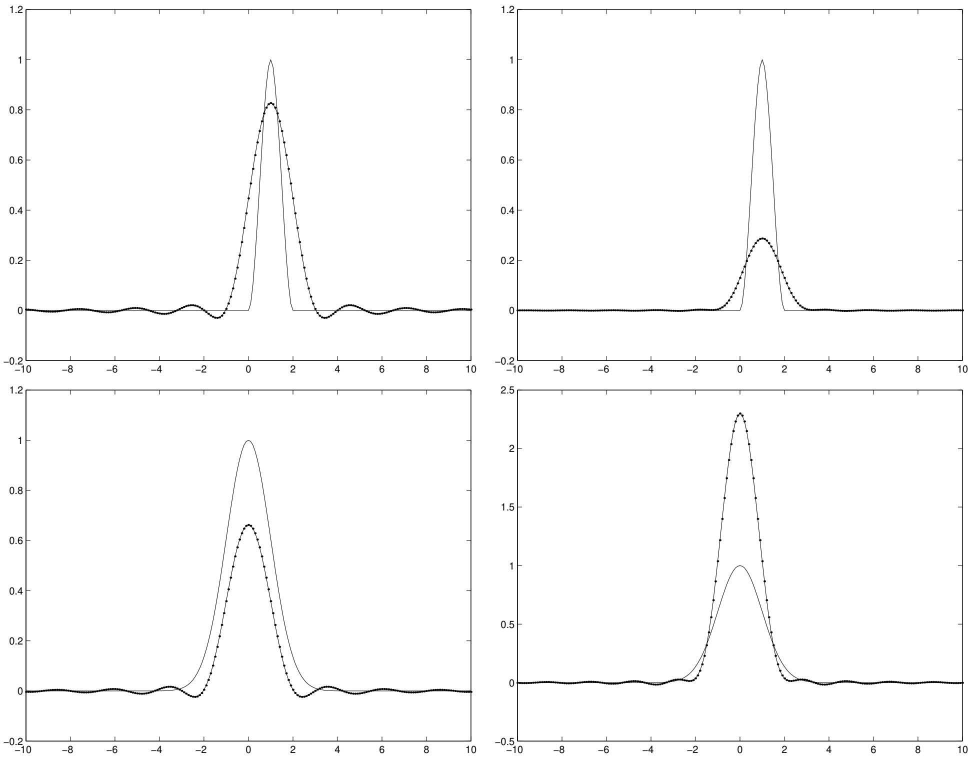

FIG. 6.5. *Reconstruction of the functions from Figure 6.1 for $\alpha = 0.01$ with overestimated (left) and underestimated γ (right).*

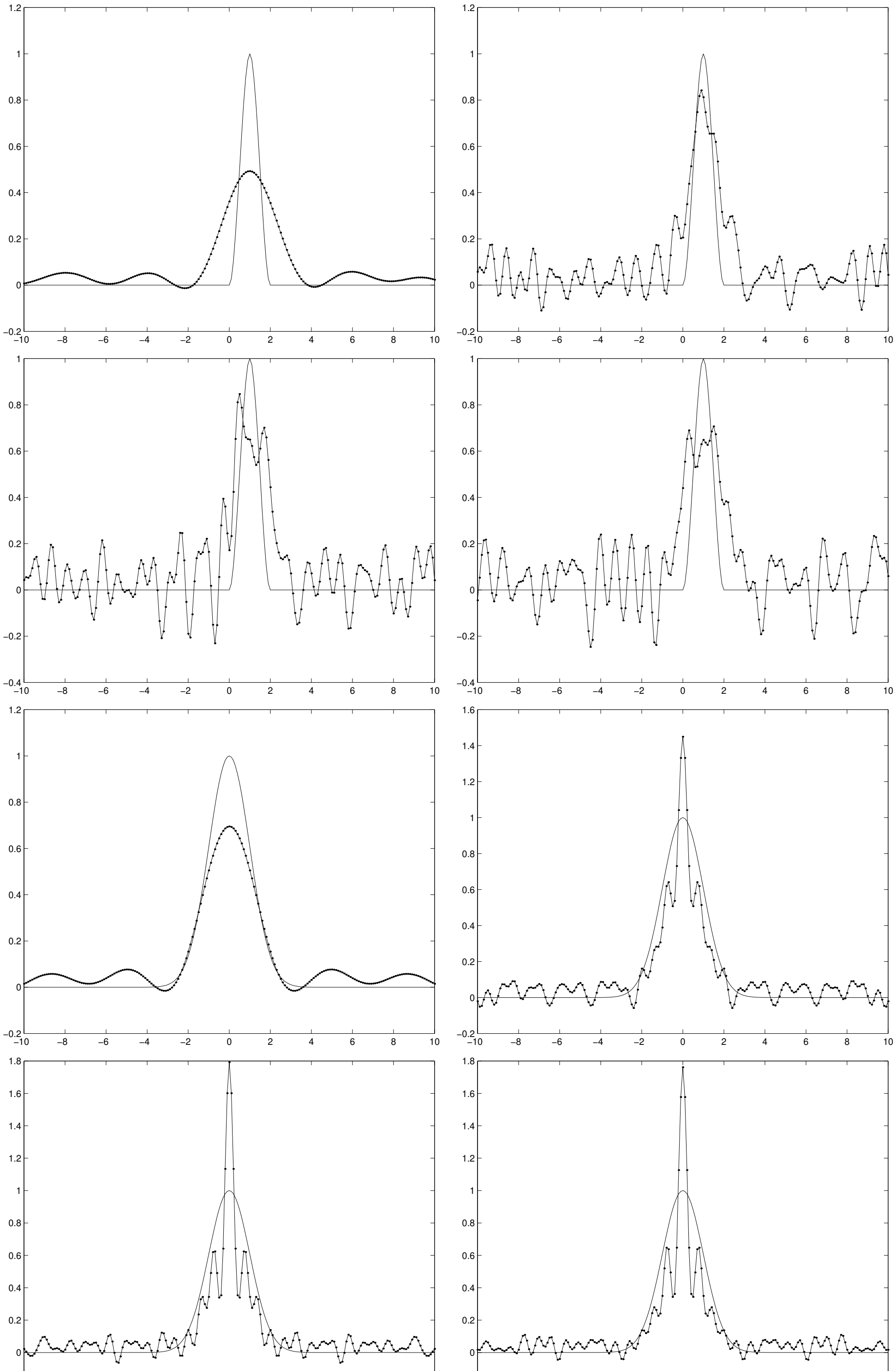

FIG. 6.6. *Reconstruction of a piecewise cubic spline image (first four pictures) and a gaussian kernel (pictures below) for the parameter values* $\alpha = 0.1, 0.01, 0.001, 0.0001$ *and noise level* 5%.

Gaussian distribution itself and the results reproduce the exact solution very well, which is similar for k in all examples, except in Figure 6.4, where the exact k is chosen to be a filter in the frequency domain. But even in this case we observe from Figure 6.4, that the Fourier transforms are reproduced very well, especially the support of the filter in the frequency domain is estimated satisfyingly. For non-smooth functions different stabilizers like the total variation seminorm are expected to produce better results (cf. [7]). An important conclusion is that the choice of the stabilizing term should always depend upon the characteristics of the specific application.

Figure 6.5 shows the effect of bad parameter choice. On the left hand side γ is chosen much larger than the one determined by (2.3). A comparison with the right hand side of Figure 6.1 shows that for large γ the function k is underestimated, for small γ the approximation is far above the exact k.

Finally, the instability due to noise in the data is illustrated in Figure 6.6, where we have perturbed the input y by random noise with noise level $\delta = 5\%$. As expected, for large values of the regularization parameters, the noise produces a minor effect, the algorithm still gives very smooth approximate solutions. For decreasing α one observes that the oscillations in x_α^δ and k_α^δ are growing, which yields large deviations to the exact solution for $\alpha \to 0$. Hence, the choice of the regularization parameter α has a strong influence on the quality of the approximations.

Acknowledgements. We thank Prof. H.W. Engl for useful and stimulating discussions.

REFERENCES

[1] R. Acar, C.R. Vogel. Analysis of total variation penalty methods. *Inverse Probl.*, 10:1217-1229, 1993.

[2] S.I. Amari. Superefficiency in Blind Source Separation *IEEE Transactions on Signal Processing*, 47:936-944, 1999.

[3] D.S.C. Biggs, and M. Andrews. Iterative blind deconvolution of extended objects. *IEEE ICIP*, 2:454-457, 1997.

[4] M. Bertero, D. Bindi, P. Bocacci, M. Cattaneo, C. Eva, and V. Lanza. A novel blind-deconvolution method with an application to seismology. *Inverse Probl.*, 14:815–833, 1998.

[5] D.C. Champerey. *A Handbook of Fourier Theorems.* Cambridge Univ. Press, Cambridge, 1987.

[6] G. Chavent, K. Kunisch. Regularization of linear least squares problems by total bounded variation. *ESAIM, Control Optim. Calc. Var.* **2**, 359-376, 1997.

[7] T.F. Chan and C.K. Wong. Total variation blind deconvolution, CAM Report 96-45, UCLA, 1996.

[8] V.C.J. Cogswell, J.A. Conchello, J.M. Lerner, T. Lu, and T.Wilson, editors. *Three-Dimensional and Multidimensional Microscopy: Image Aquisition and Processing.* Proc. SPIE, Vol. 3261, 1998.

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

[10] H.W. Engl, K. Kunisch, and A. Neubauer. Convergence rates for Tikhonov regularization of nonlinear ill–posed problems. *Inverse Probl.*, 5:523–540, 1989.

[11] E. Giusti. Minimal Surfaces and Functions of Bounded Variation. Birkhäuser, Boston, 1984.

[12] G. Hämmerlin and K.-H. Hoffmann, editors. *Constructive Methods for the Practical Treatment of Integral Equations.* Birkhäuser Verlag, Basel, Boston, Stuttgart, 1985. Proceedings of the Conference at the Mathematisches Forschungsinstitut Oberwolfach, June 24-30, 1983. International Series of Numerical Mathematics, Vol. 73.

[13] D. Kundur and D. Hatzinakos. Blind image deconvolution. *IEEE Signal Processing Magazine*, 13:43-64, 1996.

[14] D. Kundur and D. Hatzinakos. Blind image deconvolution revisted. *IEEE Signal Processing Magazine*, 13:61–63, 1996.

[15] J.L. Lions, E. Magenes. Non–Homogeneous Boundary Value Problems and Applications I–III. Springer–Verlag, New York, 1972.

[16] J. Markham, J.A. Conchello. Parametric blind deconvolution of microscopic images: further results, In *[8]*, pages 38–49, 1998.

[17] E. Schock. Approximate solution of ill-posed equations: Arbitrarily slow convergence vs. superconvergence. In *[12]*, pages 234–243, 1985.

[18] T.I. Seidman and C.R. Vogel. Well posedness and convergence of some regularisation methods for non–linear ill posed problems. *Inverse Probl.*, 5:227–238, 1989.

[19] C. Vogel, T. Chan, R. Plemmons. Fast algorithms for phase diversity-based blind deconvolution. CAM Report 98-22, UCLA, 1998.

[20] J. Weidmann. *Lineare Operatoren in Hilberträumen.* Teubner, Stuttgart, 1976.

[21] J. Wloka. *Partielle Differentialgleichungen. Sobolevräume und Randwertaufgaben.* Teubner, Stuttgart, 1982.