A Convergence Rates Result for Tikhonov Regularization in Banach Spaces with Non-Smooth Operators

B. Hofmann [‡], B. Kaltenbacher [†], C. Pöschl[*] and O. Scherzer[*]

May 28, 2008

Abstract

There exists a vast literature on convergence rates results for Tikhonov regularized minimizers. We are concerned with the solution of nonlinear ill-posed operator equations. The first convergence rates results for such problems have been developed by Engl, Kunisch and Neubauer in 1989. While these results apply for operator equations formulated in Hilbert spaces, the results of Burger and Osher from 2004, more generally, apply to operators formulated in Banach spaces. Recently, Resmerita et al. presented a modification of the convergence rates result of Burger and Osher which turns out a complete generalization of the rates result of Engl et. al. In all these papers relatively strong regularity assumptions are made. However, it has been observed numerically, that violations of the smoothness assumptions of the operator do not necessarily affect the convergence rate negatively. We take this observation and weaken the smoothness assumptions on the operator and prove a novel convergence rate result. The most significant difference in this result to the previous ones is that the source condition is formulated as a variational inequality and not as an equation as before. As examples we present a phase retrieval problem and a specific inverse option pricing problem, both studied in the literature before. For the inverse finance problem, the new approach allows us to bridge the gap to a singular case, where the operator smoothness degenerates just when the degree of ill-posedness is minimal.

[*]Department of Computer Science, University of Innsbruck, Technikerstraße 21a, 6020, Innsbruck, Austria (Email: Christiane.Poeschl,Otmar.Scherzer@uibk.ac.at). The work of OS is supported by the Austrian Science Foundation (FWF) Projects Y-123INF, FSP 9203-N12 and FSP 9207-N12. The work of CP is supported by DOC-FForte of the Austrian Academy of Sciences.

[†]Department of Mathematics, University of Stuttgart, Pfaffenwaldring 57, 70569 Stuttgart, Germany (Email: Barbara.Kaltenbacher@mathematik.uni-stuttgart.de)

[‡]Department of Mathematics, Chemnitz University of Technology, Reichenhainer Str. 41, 09107 Chemnitz, Germany (Email: hofmannb@mathematik.tu-chemnitz.de), supported by Deutsche Forschungsgemeinschaft under DFG-grant HO1454/7-1.

1 Introduction

In this paper we study variational methods for the solution of *inverse* and *ill–posed* problems, which can be written in a Banach space setting in form of an operator equation

$$F(u) = v\ . \tag{1}$$

We assume that only noisy data v^δ of the exact data v are available.

Tikhonov suggested (see for instance the book of Morozov [37]) to use minimizers of the functional

$$\boxed{\mathcal{T}_\alpha(u) := (\mathrm{dist}(F(u), v^\delta))^2 + \alpha\mathcal{R}(u)\ .}$$

for the stable approximation of solutions of (1), where $\mathrm{dist}(\cdot,\cdot)$ denotes some distance function measuring deviations in the data space. In this paper we consider a particular instance of such variational regularization models consisting in minimization of

$$\boxed{\mathcal{T}_\alpha(u) := \left\|F(u) - v^\delta\right\|_V^p + \alpha\mathcal{R}(u)\,,} \tag{2}$$

where $F : \mathcal{D}(F) \subseteq U \to V$ is the (in general nonlinear) forward operator mapping between Banach spaces U and V and where we have $1 \le p < \infty$ for the exponent in (2). Moreover, $\mathcal{R} : U \to [0, +\infty]$ is a convex and proper *stabilizing functional* with domain

$$\mathcal{D}(\mathcal{R}) := \{u \in U : \mathcal{R}(u) \neq +\infty\}\,.$$

We recall that $\mathcal{R}$ is called *proper* if $\mathcal{D}(\mathcal{R}) \neq \emptyset$.

This work has several objectives:

1. In the standard theory of variational regularization methods it is assumed that F is smooth, i.e., Fréchet derivatives exist and are also smooth (see e.g. [21]). However, controversial to the literature on regularization methods, it is often the case in applications that singularities (or nonsmooth parts) in the solution, resulting from nonsmooth parts of F, can be recovered efficiently. This motivates to develop an analysis of convergence rates of variational regularization method with nonsmooth operators. An application of such a convergence rates result with nonsmooth operators to phase retrieval is given.

2. On the other hand there exist inverse problems, for example specific inverse problems of option pricing in finance, where the smoothness of the forward operator F is determined by configurations of external model parameters. Then situations with degenerating or nonsmooth derivatives of F can just coincide with situations where the degree of ill-posedness is essentially smaller than in situations with smooth derivatives and hence the chances of good reconstruction are much better. This is the case for at-the-money options when a time-dependent local volatility function is recovered from option prices depending on varying maturities and some fixed strike price (see [26], [27]).

3. The paper is a generalization of convergence rates results of nonlinear ill–posed problem which have subsequently been proven and generalized starting from Engl, Kunisch, Neubauer [22], Burger, Osher [9], Resmerita et al [38].

2 Notation and Assumptions

Whenever this is appropriate, we omit the subscripts in the norms, dual pairings and under convergence symbols. The spaces, topologies, and notions of convergence can be identified from the context.

In this section we make the following assumptions:

Assumption 2.1. *1. U and V are Banach spaces, with which there are associated topologies τ_U and τ_V, which are weaker than the norm topologies.*

2. $\|\cdot\|_V$ is sequentially lower semi-continuous with respect to τ_V, i.e. for $v_k \to v$ with respect to the τ_V topology

$$\|v\|_V \leq \liminf_{k\to\infty} \|v_k\|_V \ .$$

3. $F : \mathcal{D}(F) \subset U \to V$ is continuous with respect to the topologies τ_U and τ_V.

4. $\mathcal{R} : U \to [0, +\infty]$ is proper, convex and τ_U lower semi continuous.

5. $\mathcal{D}(F)$ is closed with respect to τ_U and $\mathcal{D} := \mathcal{D}(F) \cap \mathcal{D}(\mathcal{R}) \neq \emptyset$.

6. For every $\alpha > 0$ and $M > 0$ the sets

$$\mathcal{M}_\alpha(M) := \{u \in \mathcal{D} : \mathcal{T}_\alpha(u) \leq M\} \ , \tag{3}$$

are τ_U sequentially compact in the following sense: every sequence (u_k) in $\mathcal{M}_\alpha(M)$ has a subsequence, which is convergent in U with respect to the τ_U-topology.

A–priori we do not exclude the case $\mathcal{M}_\alpha(M) = \emptyset$.

Remark 2.2. *Typical examples of weaker topologies τ_U, τ_V are the* weak topologies *(respectively* weak* *topologies) on Banach spaces U and V. Let $M > 0$, then the sets $\mathcal{M}_\alpha(M)$ are inversely ordered. That is*

$$\mathcal{M}_\alpha(M) \subseteq \mathcal{M}_\beta(M) \,, \quad 0 < \beta \leq \alpha \,.$$

The stabilizing character of the functional $\mathcal{R}(\cdot)$ corresponds with the fact that requirement 6 of Assumption 2.1 is satisfied.

In the Banach space theory of variational regularization methods the *Bregman distance* plays an important role.

Definition 2.3. *Let $\mathcal{R} : U \to [0, +\infty]$ be a convex and proper functional with subdifferential $\partial\mathcal{R}$. The* Bregman distance *of $\mathcal{R}$ at $u \in U$ and $\xi \in \partial\mathcal{R}(u) \subseteq U^*$ is defined by*

$$D_\xi(\tilde{u}, u) := \mathcal{R}(\tilde{u}) - \mathcal{R}(u) - \langle \xi, \tilde{u} - u \rangle \ , \quad u, \tilde{u} \in U \ .$$

Here $\langle \cdot, \cdot \rangle$ denotes the dual pairing with respect to U^* and U.

The Bregman distance is only defined at a point $\tilde{u} \in \mathcal{D}(\mathcal{R})$ where the subgradient is not empty. Moreover, the Bregman distance can be $+\infty$. The set

$$\mathcal{D}_B(\mathcal{R}) := \{u \in \mathcal{D}(\mathcal{R}) : \partial\mathcal{R}(u) \neq \emptyset\} \ .$$

is called the *Bregman domain.*

We recall that if $\mathcal{R}(u) = \|u - u_0\|_U^2$ in a Hilbert space, then $D_\xi(\tilde{u}, u) = \|\tilde{u} - u\|^2$.

3 Well-posedness

In this section we prove well-posedness, stability, and convergence of variational regularization methods consisting in minimization of (2).

Theorem 3.1. (Well-Posedness) *Assume that $\alpha > 0$, $v^\delta \in V$. Let F, $\mathcal{R}$, $\mathcal{D}$, U and V satisfy Assumption 2.1. Then there exists a minimizer of $\mathcal{T}_\alpha$.*

Proof. Since $\mathcal{D} \neq \emptyset$ and $v^\delta \in V$ there exists at least one $\tilde{u} \in U$ such that $\mathcal{T}_\alpha(\tilde{u}) < \infty$. There is a sequence (u_k) in $\mathcal{D}$ such that for $c = \inf \{\mathcal{T}_\alpha(u) : u \in \mathcal{D}\}$

$$\lim_{k \to \infty} \mathcal{T}_\alpha(u_k) = c \, .$$

From Assumption 2.1 it follows that (u_k) has a τ_U convergent subsequence which we denote again by (u_k) and the associated limit is denoted by $\tilde{u}$. Moreover, since $\mathcal{R}$ is lower semi continuous with respect to the τ_U-topology we have

$$\mathcal{R}(\tilde{u}) \leq \liminf_{k \to \infty} \mathcal{R}(u_k) \, . \tag{4}$$

By assumption, F is continuous with respect to the topologies τ_U and τ_V on $\mathcal{D}$, which is τ_U closed, showing that $\tilde{u} \in \mathcal{D}$. Therefore, $F(u_k) - v^\delta$ converges to $F(\tilde{u}) - v^\delta$ with respect to τ_V.

Since $\|\cdot\|_V$ is sequentially lower continuous with respect to the τ_V-topology it follows that

$$\left\|F(\tilde{u}) - v^\delta\right\|_V^p \leq \liminf_{k \to \infty} \left\|F(u_k) - v^\delta\right\|_V^p \, . \tag{5}$$

Combination of (4) and (5) shows that $\tilde{u}$ minimizes $\mathcal{T}_\alpha$. □

Theorem 3.2. (Stability) *The minimizers of $\mathcal{T}_\alpha$ are stable with respect to the data v^δ. That is, if (v_k) is a sequence converging to v^δ in V with respect to the norm-topology, then every sequence (u_k) satisfying*

$$u_k \in \mathit{argmin}\{\|F(u) - v_k\|_V^p + \alpha\mathcal{R}(u) : u \in \mathcal{D}\} \tag{6}$$

has a subsequence, which converges with respect to the τ_U topology, and the limit of each τ_U convergent subsequence is a minimizer $\tilde{u}$ of $\mathcal{T}_\alpha$ as in (2).

Moreover, for each τ_U convergent subsequence (u_m), $(\mathcal{R}(u_m))$ converges to $\mathcal{R}(\tilde{u})$.

Proof. From the definition of u_k it follows that

$$\|F(u_k) - v_k\|_V^p + \alpha\mathcal{R}(u_k) \le \|F(u) - v_k\|_V^p + \alpha\mathcal{R}(u)\,, \quad u \in \mathcal{D}\,.$$

Since $\mathcal{D} \neq \emptyset$ we can select $\overline{u} \in \mathcal{D}$ and since $v_k \to v^\delta$ with respect to the norm topology, it follows that

$$\begin{aligned}
&\|F(u_k) - v\|_V^p + \alpha\mathcal{R}(u_k) \\
\le &2^{p-1}\|F(u_k) - v_k\|_V^p + 2^{p-1}\alpha\mathcal{R}(u_k) + 2^{p-1}\|v_k - v\|_V^p \\
\le &2^{p-1}\|F(\overline{u}) - v_k\|_V^p + 2^{p-1}\alpha\mathcal{R}(\overline{u}) + 2^{p-1}\|v_k - v\|_V^p\,.
\end{aligned}$$

Thus for every $\epsilon > 0$ there exists $k_0 \in \mathbb{N}$ such that for $k \ge k_0$

$$\|F(u_k) - v\|_V^p + \alpha\mathcal{R}(u_k) \le 2^{p-1}\|F(\overline{u}) - v\|_V^p + \alpha\mathcal{R}(\overline{u})\epsilon = M\,.$$

Thus (u_k) is in $\mathcal{M}_\alpha(M)$ and therefore has a τ_U convergent subsequence in U. Now, let (u_k) denote an arbitrary τ_U convergent subsequence with limit $\tilde{u} \in \mathcal{D}$. Since F is continuous with respect to the τ_U and τ_V topologies it follows that $F(u_k) \to F(\tilde{u})$ with respect to the τ_V topology. Moreover, since the τ_V topology is weaker than the norm topology it follows that $v_k \to v^\delta$ with respect to the τ_V topology and thus $F(u_k) - v_k$ converges to $F(\tilde{u}) - v^\delta$ with respect to the τ_V topology. Since $\|\cdot\|_V$ and $\mathcal{R}(\cdot)$ are lower semi continuous with respect to the τ_V and τ_U topology, respectively, it follows that

$$\begin{aligned}
\left\|F(\tilde{u}) - v^\delta\right\|_V^p &\le \liminf_{k\to\infty} \|F(u_k) - v_k\|_V^p\,, \\
\mathcal{R}(\tilde{u}) &\le \liminf_{k\to\infty} \mathcal{R}(u_k)\,.
\end{aligned} \tag{7}$$

Using the results above it follows that

$$\begin{aligned}
\left\|F(\tilde{u}) - v^\delta\right\|_V^p + \alpha\mathcal{R}(\tilde{u}) &\le \liminf_{k\to\infty} \|F(u_k) - v_k\|_V^p + \alpha \liminf_{k\to\infty} \mathcal{R}(u_k) \\
&\le \limsup_{k\to\infty} \left(\|F(u_k) - v_k\|_V^p + \alpha\mathcal{R}(u_k)\right) \\
&\le \lim_{k\to\infty} \left(\|F(u) - v_k\|_V^p + \alpha\mathcal{R}(u)\right) \\
&= \left\|F(u) - v^\delta\right\|_V^p + \alpha\mathcal{R}(u)\,, \quad u \in \mathcal{D}\,.
\end{aligned}$$

This implies that $\tilde{u}$ is a minimizer and moreover by taking $u = \tilde{u} \in \mathcal{D}$ on the right hand side it follows that

$$\left\|F(\tilde{u}) - v^\delta\right\|_V^p + \alpha\mathcal{R}(\tilde{u}) = \lim_{k\to\infty} \left(\|F(u_k) - v_k\|_V^p + \alpha\mathcal{R}(u_k)\right)\,. \tag{8}$$

Now assume that $\mathcal{R}(u_k)$ does not converge to $\mathcal{R}(\tilde{u})$. Since $\mathcal{R}$ is lower semi continuous with respect to the τ_U topology it follows then that

$$c := \limsup_{k\to\infty} \mathcal{R}(u_k) > \mathcal{R}(\tilde{u}) .$$

We take a subsequence (u_k) such that $\mathcal{R}(u_k) \to c$. For this subsequence we find as a consequence of (8) that

$$\lim_{k\to\infty} \|F(u_k) - v_k\|_V^p = \left\|F(\tilde{u}) - v^\delta\right\|_V^p + \alpha\left(\mathcal{R}(\tilde{u}) - c\right) < \left\|F(\tilde{u}) - v^\delta\right\|_V^p .$$

This contradicts (7). Therefore we obtain $\mathcal{R}(u_k) \to \mathcal{R}(\tilde{u})$. $\square$

Assumption 2.1 is, e.g., satisfied if we take the weak topologies on U and V, for τ_U and τ_V and F is continuous with respect to the weak topologies. In the Hilbert space setting we deduce from Theorem 3.2 that a subsequence of u_k converges weakly to $\tilde{u}$ in U and that $\mathcal{R}(u_k) \to \mathcal{R}(\tilde{u})$. In the Hilbert space setting this gives strong convergence along a subsequence.

In the following we prove convergence, convergence rates, and stability estimates for variational regularization methods in Banach spaces.

The generalized solution concept in a Banach space setting is:

Definition 3.3. *An element $u^\dagger \in \mathcal{D}$ is called an $\mathcal{R}$-minimizing solution if*

$$\boxed{\mathcal{R}(u^\dagger) = \min\left\{\mathcal{R}(u) : F(u) = v\right\} < \infty .}$$

This solution concept generalizes the definition of an u_0-minimal norm solution in a Hilbert space setting.

Theorem 3.4. (Existence) *Let Assumption 2.1 be satisfied. If there exists a solution of (1), then there exists a $\mathcal{R}$-minimizing solution.*

Proof. All along this proof we consider the $\mathcal{T}_\alpha$ with v^δ replaced by v.

Suppose there does not exist an $\mathcal{R}$-minimizing solution in $\mathcal{D}$. Then there exists a sequence (u_k) of solutions of (1) in $\mathcal{D}$ such that $\mathcal{R}(u_k) \to c$ and

$$c < \mathcal{R}(u) \text{ for all } u \in \mathcal{D} \text{ satisfying } F(u) = v. \tag{9}$$

Thus for sufficiently large k and $\alpha = 1$ it follows that

$$\mathcal{T}_\alpha(u_k) = \mathcal{R}(u_k) < 2c.$$

Thus $(u_k) \subseteq \mathcal{M}_1(2c)$, and from (3) it follows that (u_k) is τ_U sequentially compact, and consequently has a τ_U convergent subsequence, which we again denote by (u_k). The τ_U-limit is denoted by $\tilde{u}$. From the τ_U lower semi continuity of $\mathcal{R}$ it follows that $\mathcal{R}(\tilde{u}) \le \liminf_{k\to\infty} \mathcal{R}(u_k) = c$.

Moreover, since F is continuous with respect to the topologies τ_U and τ_V it follows from $F(u_k) = v$ that $F(\tilde{u}) = v$. This gives a contradiction to (9). $\square$

Theorem 3.5. (Convergence) *Let F, $\mathcal{R}$, $\mathcal{D}$, U and V satisfy Assumption 2.1. Moreover, we assume that there exists a solution of (2) (Then, according to Theorem 3.4 there exists an $\mathcal{R}$-minimizing solution).*

Assume that the sequence (δ_k) converges monotonically to 0 and $v_k := v^{\delta_k}$ satisfies $\|v - v_k\|_V \leq \delta_k$.

Moreover, assume that $\alpha(\delta)$ satisfies

$$\alpha(\delta) \to 0 \text{ and } \frac{\delta^p}{\alpha(\delta)} \to 0 \text{ as } \delta \to 0 \,.$$

and $\alpha(\cdot)$ is monotonically increasing. We denote by $\alpha_k = \alpha(\delta_k)$ and with $\alpha_1 = \alpha_{\max}$.

A sequence (u_k) satisfying (6) has a convergent subsequence with respect to the τ_U-topology. A limit of each τ_U convergent subsequence is an $\mathcal{R}$-minimizing solution.

If in addition the $\mathcal{R}$-minimizing solution $u^\dagger$ is unique, then $u_k \to u^\dagger$ with respect to τ_U.

Proof. From the definition of u_k it follows that

$$\|F(u_k) - v_k\|_V^p + \alpha_k \mathcal{R}(u_k) \leq \delta_k^p + \alpha_k \mathcal{R}(u^\dagger)$$

which shows that

$$\lim_{k\to\infty} F(u_k) = v \text{ with respect to the norm topology on } V$$

and that

$$\limsup_{k\to\infty} \mathcal{R}(u_k) \leq \mathcal{R}(u^\dagger) \,. \tag{10}$$

Therefore, we have

$$\begin{aligned}
&\limsup_{k\to\infty} \left(\|F(u_k) - v_k\|_V^p + \alpha_{\max}\mathcal{R}(u_k)\right) \\
\leq &\limsup_{k\to\infty} \left(\|F(u_k) - v_k\|_V^p + \alpha_k\mathcal{R}(u_k)\right) + \limsup_{k\to\infty} (\alpha_{\max} - \alpha_k)\mathcal{R}(u_k) \\
&\leq \alpha_{\max}\mathcal{R}(u^\dagger) < \infty \,.
\end{aligned}$$

From Assumption 2.1 it follows that (u_k) has a subsequence, which is again denoted by (u_k), which converges with respect to the τ_U topology to some $\tilde{u} \in \mathcal{D}$. Using that F is continuous with respect to the topology τ_V, and that the norm convergence on V is stronger, it follows from (10) that $F(\tilde{u}) = v$.

From the lower semi continuity of $\mathcal{R}$ with respect to the τ_U topology it follows that

$$\mathcal{R}(\tilde{u}) \leq \liminf_{k\to\infty} \mathcal{R}(u_k) \leq \limsup_{k\to\infty} \mathcal{R}(u_k) \leq \mathcal{R}(u^\dagger) \leq \mathcal{R}(u_s) \,, \tag{11}$$

for all $u_s \in \mathcal{D}$ satisfying $F(u_s) = v$. Taking $u_s = \tilde{u}$ shows that $\mathcal{R}(\tilde{u}) = \mathcal{R}(u^\dagger)$. That is $\tilde{u}$ is an $\mathcal{R}$-minimizing solution.

Using this and (11) it follows that $\mathcal{R}(u_k) \to \mathcal{R}(u^\dagger)$.

If the $\mathcal{R}$-minimizing solution is unique it follows that (u_k) has a τ_U-convergent subsequence and the limit of any τ_U-convergent subsequence of (u_k) has to be equal to $u^\dagger$. Therefore, a subsequence-subsequence argument implies convergence of the whole sequence. □

Remark 3.6. *Given $\alpha_{\max} > 0$ fixed.*

By u_α^δ we denote a minimizer of the functional (2)*. Under the assumption of Theorem 3.5 it follows that for sufficiently small δ, $\alpha(\delta) \le \alpha_{\max}$ and therefore*

$$\mathcal{R}(u_\alpha^\delta) \le \frac{\delta^p}{\alpha} + \mathcal{R}(u^\dagger)$$

and

$$\begin{aligned}
&\left\|F(u_\alpha^\delta) - v^\delta\right\|_V^p + \alpha_{\max}\mathcal{R}(u_\alpha^\delta)\\
\le&\left\|F(u_\alpha^\delta) - v^\delta\right\|_V^p + \alpha\mathcal{R}(u_\alpha^\delta) + (\alpha_{\max} - \alpha)\mathcal{R}(u_\alpha^\delta)\\
\le&\left\|F(u^\dagger) - v^\delta\right\|_V^p + \alpha\mathcal{R}(u^\dagger) + (\alpha_{\max} - \alpha)\mathcal{R}(u_\alpha^\delta)\\
\le&\alpha_{\max}\left(\mathcal{R}(u^\dagger) + \frac{\delta^p}{\alpha}\right) .
\end{aligned}$$

This shows that

$$u_\alpha^\delta \in \mathcal{M}_{\alpha_{\max}}\left(\alpha_{\max}\left(\mathcal{R}(u^\dagger) + \frac{\delta^p}{\alpha}\right)\right) . \tag{12}$$

4 The Convergence Rates Result

To show convergence rates we need to make the following assumptions:

Assumption 4.1. *F, $\mathcal{R}$, U, V and $\mathcal{D}$ satisfy Assumption 2.1.*

1. *There exists an $\mathcal{R}$-minimizing solution $u^\dagger$ which is an element of the Bregman domain $\mathcal{D}_B(\mathcal{R})$.*
2. *Let $\rho > \alpha_{max}\left(\mathcal{R}(u^\dagger) + \frac{\delta^p}{\alpha}\right)$. It follows from (12) that $u_\alpha^\delta, u^\dagger \in \mathcal{M}_\alpha(\rho)$.*
3. *v^δ satisfies*
$$\left\|v - v^\delta\right\|_V \le \delta . \tag{13}$$
4. *There exist numbers $\beta_1, \beta_2 \in [0,\infty)$, with $\beta_1 < 1$, and $\xi \in \partial\mathcal{R}(u^\dagger)$ such that*
$$\left\langle \xi, u^\dagger - u\right\rangle_{U^*,U} \le \beta_1 D_\xi(u, u^\dagger) + \beta_2\left\|F(u) - F(u^\dagger)\right\|_V \tag{14}$$
for all $u \in \mathcal{M}_{\alpha_{\max}}(\rho)$.

Remark 4.2. *If*

- *$V \subseteq \tilde{V}$, such that*
$$\langle \tilde{v}^*, v\rangle_{V^*,V} = \langle \tilde{v}^*, v\rangle_{\tilde{V}^*,\tilde{V}} , \quad \tilde{v}^* \in \tilde{V}^*, v \in V , \tag{15}$$

- *$\mathcal{D}$ is starlike with respect to $u^\dagger$, that is, for every $u \in \mathcal{D}$ there exists t_0 such that*

$$u^\dagger + t\,(u^\dagger - u) \in \mathcal{D} \qquad \forall t \leq t_0\,,$$

- *$F : \mathcal{D} \to V$ attains a one-sided directional derivative at $u^\dagger$, that is, for every $u \in \mathcal{D}$ the element*

$$\lim_{t \to 0+} \frac{1}{t}\left(F(u^\dagger + t(u^\dagger - u)) - F(u^\dagger)\right) = F'(u^\dagger; u^\dagger - u) \in V$$

  exists,

then 4. in Assumption 4.1 holds if

1. *there exists $\gamma > 0$ such that*

$$\left\|F(u) - F(u^\dagger) - F'(u^\dagger; u - u^\dagger)\right\|_{\tilde{V}} \leq \gamma D_\xi(u, u^\dagger)\,, \quad u \in \mathcal{M}_{\alpha_{max}}(\rho) \quad (16)$$

   and

2. *there exist $\tilde{\omega} \in \tilde{V}^*$ and $\xi \in \partial\mathcal{R}(u^\dagger)$ such that for all $u \in \mathcal{M}_{\alpha_{max}}(\rho)$*

$$\left\langle \xi, u^\dagger - u\right\rangle_{U^*,U} \leq \left|\left\langle \tilde{\omega}, F'(u^\dagger; u - u^\dagger)\right\rangle_{\tilde{V}^*,\tilde{V}}\right| \text{ and } \gamma\|\tilde{\omega}\|_{\tilde{V}^*} < 1. \quad (17)$$

This can be seen by choosing $\beta_1 = \gamma\|\tilde{\omega}\|_{\tilde{V}^*}$ and $\beta_2 = \|\tilde{\omega}\|_{V^*}$ such that we have

$$\begin{aligned}
&\left|\left\langle \tilde{\omega}, F'(u^\dagger; u - u^\dagger)\right\rangle_{\tilde{V}^*,\tilde{V}}\right| \\
\leq& \|\tilde{\omega}\|_{V^*}\left\|F(u) - F(u^\dagger)\right\|_V + \|\tilde{\omega}\|_{\tilde{V}^*}\left\|F(u) - F(u^\dagger) - F'(u^\dagger; u - u^\dagger)\right\|_{\tilde{V}} \\
\leq& \|\tilde{\omega}\|_{V^*}\left\|F(u) - F(u^\dagger)\right\|_V + \gamma\|\tilde{\omega}\|_{\tilde{V}^*} D_\xi(u, u^\dagger)\,.
\end{aligned}$$

We also highlight (15). In the case $V := L^2(0,1) \subset \tilde{V} := L^{2-\epsilon}(0,1)$ with $0 < \epsilon < 1$ (used later) we have that V^* and $\tilde{V}^*$ are isomorph to $L^2(0,1)$ and $L^{(2-\epsilon)^*}(0,1)$, with isomorphisms i_1, i_2, respectively. Thus we can write

$$\langle \tilde{v}^*, v\rangle_{V^*,V} = \int_0^1 i_1(\tilde{v}^*)v = \int_0^1 i_2(\tilde{v}^*)v = \langle \tilde{v}^*, v\rangle_{\tilde{V}^*,\tilde{V}}\,,$$

which is equivalent to $i_1(\tilde{v}^*) = i_2(\tilde{v}^*)$.

Remark 4.3. *If $V = \tilde{V}$ and F is Gâteaux differentiable,*

$$F'(u^\dagger)^* : V^* \to U^*,$$

denotes the adjoint operator of $F'(u^\dagger)$ which is defined by

$$\left\langle F'(u^\dagger)^* v^*, u\right\rangle_{U^*,U} = \left\langle v^*, F'(u^\dagger)u\right\rangle_{V^*,V}\,, \quad u \in U,\ v^* \in V^*\,.$$

Under these particular assumptions and the notation $\omega = \tilde{\omega}$ (17) holds if

$$\left\langle \xi, u^\dagger - u\right\rangle_{U^*,U} \leq \left|\left\langle F'(u^\dagger)^*\omega, u - u^\dagger\right\rangle_{U^*,U}\right| \text{ and } \gamma\|\omega\|_{V^*} < 1 \quad (18)$$

for all $u \in \mathcal{M}_{\alpha_{max}}(\rho)$.

In the special situation of classical convergence rates (cf. as seminal paper [22]), where $V = \tilde{V}$ *and* U *are Hilbert spaces and* F *is Fréchet-differentiable,* (18) *is equivalent to saying*

$$\text{there exists} \quad \overline{\omega} \text{ with } \gamma\|\overline{\omega}\|_{V^*} < 1 \text{ such that } \xi = F'(u^\dagger)^*\overline{\omega}\ . \tag{19}$$

Thus (18) is a generalization of the standard source condition (sourcewise representation) of the solution in convergence rates results for the Tikhonov regularization (cf. [21, Chapter 10]).

However, note that (18) is in general a nonlinear *condition.*

- (19) $\Rightarrow$ (18). *If* (19) *holds then*

  $$\left\langle \xi, u - u^\dagger \right\rangle_{U^*,U} = \left\langle F'(u^\dagger)^*\overline{\omega}, u - u^\dagger \right\rangle_{U^*,U} \text{ and } \gamma\|\overline{\omega}\|_{V^*} < 1\ ,$$

  thus (18) *holds for* $\omega = \overline{\omega}$.

- (18) $\Rightarrow$ (19). *Let us assume that there exists a singular value decomposition* $(u_k, v_k, \sigma_k)_{k\in\mathbb{N}}$ *of* $F'(u^\dagger)$. *That is,* $\left\{(\sigma_k)^2 : k \in \mathbb{N}\right\}$ *are the non-zero eigenvalues of the operator* $F'(u^\dagger)^*F'(u^\dagger)$, *written down in decreasing order with multiplicity, and* $\{u_k : k \in \mathbb{N}\}$ *the set of corresponding eigenvectors which span the closure of the range of* $F'(u^\dagger)^*F'(u^\dagger)$. *Moreover, let* $v_k = \frac{F'(u^\dagger)u_k}{\|F'(u^\dagger)u_k\|}$. *For a spectral decomposition, the following holds*

  $$\begin{aligned} F'(u^\dagger)u_k &= \sigma_k v_k\ , \quad F'(u^\dagger)u = \sum_{k=1}^{\infty} \sigma_k \langle u, u_k \rangle v_k\ , \quad u \in U\ , \\ F'(u^\dagger)^*v_k &= \sigma_k u_k\ , \quad F'(u^\dagger)^*v = \sum_{k=1}^{\infty} \sigma_k \langle v, v_k \rangle u_k\ , \quad v \in V\ . \end{aligned}$$

  With $\xi := \sum \xi^i u_i$, $\omega := \sum \omega^i v_i$ *and* $u := \pm u_k + u^\dagger$ (18) *reads*

  $$\begin{aligned} \pm\xi^k = \langle \xi, \pm u_k \rangle_{U^*,U} &\le \left| \left\langle F'(u^\dagger)^*\omega, \pm u_k \right\rangle_{U^*,U} \right| \\ &= \left| \sum_{l=1}^{\infty} \sigma_l \langle \omega, v_l \rangle_{V^*,V} \langle u_l, u_k \rangle_{U^*,U} \right| = \left|\sigma_k \omega^k\right| \end{aligned}$$

  and $\gamma\|\omega\|_{V^*} = \gamma\sqrt{\sum(\omega^k)^2} < 1$. *Hence we have* $\left|\xi^k\right| \le \left|\sigma_k\omega^k\right|$.
  Define $\overline{\omega} := \sum \overline{\omega}^k v_k$ *with with* $\xi^k = \sigma_k\overline{\omega}^k$. *Then*

  $$\xi = F'(u^\dagger)^*\overline{\omega} \quad \text{and } \gamma\|\overline{\omega}\|_{V^*} \le \gamma\|\omega\|_{V^*} < 1\,.$$

  Thus (19) *is fullfilled.*

Using the spectral theorem for bounded selfadjoint operators, the same can be shown for non-compact operators.

Theorem 4.4. (Convergence Rates) *Assume that F, $\mathcal{R}$, $\mathcal{D}$, U and V satisfy Assumption 4.1.*

- $p > 1$. *For $\alpha : (0, \infty) \to (0, \infty)$ satisfying $c\delta^{p-1} \leq \alpha(\delta) \leq C\delta^{p-1}$ $(0 < c \leq C)$ we have*

$$D_\xi(u_\alpha^\delta, u^\dagger) = O(\delta) \text{ and } \left\| F(u_\alpha^\delta) - v^\delta \right\|_V = O(\delta)\ .$$

  Moreover from the definition of u_α^δ it follows that

$$\mathcal{R}(u_\alpha^\delta) \leq \mathcal{R}(u^\dagger) + \frac{\delta}{c}.$$

- $p = 1$. *For $\alpha : (0, \infty) \to (0, \infty)$ satisfying $c\delta^\epsilon \leq \alpha(\delta) \leq C\delta^\epsilon$ $(0 < c \leq C, 0 < \epsilon < 1)$, we have*

$$D_\xi(u_\alpha^\delta, u^\dagger) = O(\delta^{1-\epsilon}) \text{ and } \left\| F(u_\alpha^\delta) - v^\delta \right\|_V = O(\delta)\ .$$

  Moreover from the definition of u_α^δ it follows that

$$\mathcal{R}(u_\alpha^\delta) \leq \mathcal{R}(u^\dagger) + \frac{\delta^{1-\epsilon}}{c}.$$

Proof. From the definition of u_α^δ and (13) it follows that

$$\left\| F(u_\alpha^\delta) - v^\delta \right\|_V^p + \alpha D_\xi(u_\alpha^\delta, u^\dagger) \leq \delta^p + \alpha \left(\mathcal{R}(u^\dagger) - \mathcal{R}(u_\alpha^\delta) + D_\xi(u_\alpha^\delta, u^\dagger) \right)\ . \tag{20}$$

Using (14) it follows that

$$\begin{aligned}
&\mathcal{R}(u^\dagger) - \mathcal{R}(u_\alpha^\delta) + D_\xi(u_\alpha^\delta, u^\dagger)\\
&= - \left\langle \xi, u_\alpha^\delta - u^\dagger \right\rangle_{U^*,U}\\
&\leq \beta_1\, D_\xi(u_\alpha^\delta, u^\dagger) \,+\, \beta_2 \left\| F(u_\alpha^\delta) - F(u^\dagger) \right\|_V\\
&\leq \beta_1\, D_\xi(u_\alpha^\delta, u^\dagger) \,+\, \beta_2 \left(\left\| F(u_\alpha^\delta) - v^\delta \right\|_V + \delta \right)
\end{aligned}$$

Therefore from (20) it follows that

$$\begin{aligned}
&\left\| F(u_\alpha^\delta) - v^\delta \right\|_V^p + \alpha D_\xi(u_\alpha^\delta, u^\dagger)\\
\leq\ & \delta^p + \alpha \left(\mathcal{R}(u^\dagger) - \mathcal{R}(u_\alpha^\delta) + D_\xi(u_\alpha^\delta, u^\dagger) \right)\\
\leq\ & \delta^p + \alpha \left(\beta_1 D_\xi(u_\alpha^\delta, u^\dagger) \,+\, \beta_2 \left(\left\| F(u_\alpha^\delta) - v^\delta \right\|_V + \delta \right) \right)
\end{aligned} \tag{21}$$

- **Case** $p = 1$. From (21) it follows that

$$\begin{aligned}
&(1 - \alpha\, \beta_2) \left\| F(u_\alpha^\delta) - v^\delta \right\|_V + \alpha\, (1 - \beta_1)\, D_\xi(u_\alpha^\delta, u^\dagger)\\
\leq&\delta + \alpha\delta\, \beta_2\ .
\end{aligned}$$

  This shows that

$$\left\| F(u_\alpha^\delta) - v^\delta \right\|_V \leq \delta \frac{1 + \alpha\, \beta_2}{1 - \alpha\, \beta_2} \tag{22}$$

and

$$D_\xi(u_\alpha^\delta, u^\dagger) \leq \frac{\delta\,(1+\alpha\,\beta_2)}{\alpha\,(1-\beta_1)}\ . \tag{23}$$

Taking into account the choice of $\alpha = \alpha(\delta)$ the assertion follows.

- **Case** $p > 1$**.** From (21) it follows that

$$\begin{aligned}
&\left(\left\|F(u_\alpha^\delta) - v^\delta\right\|_V^{p-1} - \alpha\beta_2\right)\left\|F(u_\alpha^\delta) - v^\delta\right\|_V \\
&+ \alpha\,(1-\beta_1)\,D_\xi(u_\alpha^\delta, u^\dagger) \\
&\leq \delta^p + \alpha\delta\,\beta_2
\end{aligned} \tag{24}$$

Using Young's inequality

$$ab \leq \frac{a^p}{p} + \frac{b^{p_*}}{p_*}, \qquad \frac{1}{p} + \frac{1}{p_*} = 1,$$

with $a = \left\|F(u_\alpha^\delta) - v^\delta\right\|_V$ and $b = \alpha\,\beta_2$

$$-\frac{1}{p}\left\|F(u_\alpha^\delta) - v^\delta\right\|_V^p \leq -\alpha\,\beta_2\left\|F(u_\alpha^\delta) - v^\delta\right\|_V + \frac{1}{p_*}\,(\alpha\,\beta_2)^{p_*}$$

it follows from (24) (taking into account that by our parameter choice $\alpha = O(\delta^{p-1})$)

$$\left\|F(u_\alpha^\delta) - v^\delta\right\|_V \leq \sqrt[p]{\frac{p}{p-1}\left(\delta^p + \alpha\delta\,\beta_2 + \frac{(\alpha\,\beta_2)^{p_*}}{p_*}\right)} = O(\delta) \tag{25}$$

and

$$D_\xi(u_\alpha^\delta, u^\dagger) \leq \frac{\delta^p + \alpha\delta\,\beta_2 + \frac{1}{p_*}\,(\alpha\,\beta_2)^{p_*}}{\alpha\,(1-\beta_1)} = O(\delta)\ . \tag{26}$$

This shows the assertion.

□

Remark 4.5. *Let $\alpha > 0$ be fixed and $\delta = 0$ (that is, we assume exact data). Let Assumption 4.1 be satisfied. Following the proof of Theorem 4.4 we see the following:*

If $p = 1$, then from (22) and (23) it follows under the assumption that fixed value α is so sufficiently small that $\alpha\beta_2 < 1$ then

$$\|F(u_\alpha) - v\|_V = 0 \text{ and } D_{\xi^*}(u_\alpha, u^\dagger) = 0\ .$$

The last identity is the reason that in [9] regularization methods with $p = 1$ are called exact penalization methods *. In the case of perturbed data, since α is fixed, it follows from (22) and (23) that*

$$\|F(u_\alpha) - v\|_V = O(\delta) \text{ and } D_{\xi^*}(u_\alpha, u^\dagger) = O(\delta)\,,$$

which is also a result stated in [9].

Let $p > 1$. From (25) and (26) it follows that

$$\|F(u_\alpha) - v\|_V \leq \sqrt[p]{\beta_2}\alpha^{1/(p-1)}$$

and

$$D_\xi(u_\alpha^\delta, u^\dagger) \leq \frac{\beta_2^{p_*}}{p_* (1-\beta_1)}\alpha^{p_*-1} .$$

Remark 4.6. *Several convergence rates results for Tikhonov regularization of the form*

$$\left\|F(u_\alpha^\delta) - v^\delta\right\|_V = O(\delta) \text{ and } D_\xi(u_\alpha^\delta, u^\dagger) = O(\delta)$$

in a Banach space setting have been derived in the literature.

1. *Chavent & Kunisch [15] have proven a convergence rates results for regularization with $\mathcal{R}(u) = \int_\Omega u^2 + |Du|$. They did not express the convergence rates result in terms of Bregman distances.*

2. *Burger & Osher [9] assumed that U is a Banach space, V is an Hilbert space, that F is Fréchet-differentiable and that there exists $\omega \in V$ and $\xi \in \partial\mathcal{R}(u^\dagger)$ (subdifferential of $\mathcal{R}$ at $u^\dagger$) which satisfies*

$$F'(u^\dagger)^*\omega = \xi \tag{27}$$

and

$$\begin{aligned} &\left\langle F(u) - F(u^\dagger) - F'(u^\dagger)(u-u^\dagger), \omega\right\rangle_V \\ \leq &\gamma\left\|F(u) - F(u^\dagger)\right\|_V \|\omega\|_V . \end{aligned} \tag{28}$$

It follows from (28) and (27) that

$$\begin{aligned} & - \left\langle \xi, u - u^\dagger\right\rangle_{U^*,U} \\ \leq & - \left\langle F'(u^\dagger)^*\omega, u - u^\dagger\right\rangle_{U^*,U} \\ = & - \left\langle \omega, F'(u^\dagger)(u-u^\dagger)\right\rangle_V \\ = & - \left\langle \omega, F'(u^\dagger)(u-u^\dagger) + F(u^\dagger) - F(u) + F(u) - F(u^\dagger)\right\rangle_V \\ \leq & \underbrace{(1+\gamma)\|\omega\|_V}_{=:\beta_2}\left\|F(u) - F(u^\dagger)\right\|_V \end{aligned}$$

Thus (14) holds. Note, that in [9] no smallness condition is associated with the source condition (27), which is not necessary since (28) is already scaling invariant.

3. *In [38] we assumed that U, V are both Banach spaces, F Fréchet-differentiable and*

$$\left\|F(u) - F(u^\dagger) - F'(u^\dagger)(u-u^\dagger)\right\|_V \leq \gamma D_\xi(u, u^\dagger) .$$

Moreover, there we assumed that there exists $\omega^* \in V^*$ satisfying

$$F'(u^\dagger)^*\omega^* = \xi \in \partial\mathcal{R}(u^\dagger) \text{ and } \gamma\|\omega^*\|_{V^*} < 1.$$

Under this assumptions, we were able to prove that assertion of Theorem 4.4 are valid.

Now, we consider a slightly more general version of [38]: we assume that $V \subset \tilde{V}$ and the generalized derivative of F satisfies

$$\left\|F(u) - F(u^\dagger) - F'(u^\dagger; u - u^\dagger)\right\|_{\tilde{V}} \leq \gamma D_\xi(u, u^\dagger) \,.$$

Moreover, we assume that there exists $\tilde{\omega}^* \in \tilde{V}^* \subset V^*$, which satisfies

$$-\left\langle \xi, u - u^\dagger \right\rangle_{U^*,U} \leq \left|\left\langle \tilde{\omega}^*, F'(u^\dagger; u - u^\dagger)\right\rangle_{\tilde{V}^*,\tilde{V}}\right| \text{ and } \gamma \|\tilde{\omega}^*\|_{\tilde{V}^*} < 1.$$

Then

$$\begin{aligned}
& -\left\langle \xi, u - u^\dagger \right\rangle_{U^*,U} \\
\leq & \left|\left\langle \tilde{\omega}^*, F'(u^\dagger; u - u^\dagger)\right\rangle_{\tilde{V}^*,\tilde{V}}\right| \\
\leq & \left|\left\langle \tilde{\omega}^*, F'(u^\dagger)(u - u^\dagger) + F(u^\dagger) - F(u) + F(u) - F(u^\dagger)\right\rangle_{\tilde{V}^*,\tilde{V}}\right| \\
\leq & \underbrace{\gamma \|\tilde{\omega}^*\|_{\tilde{V}^*}}_{=:\beta_1} D_\xi(u, u^\dagger) + \underbrace{\|\tilde{\omega}^*\|_{\tilde{V}^*}}_{=:\beta_2} \left\|F(u) - F(u^\dagger)\right\|_V.
\end{aligned}$$

Which again gives (14) if $\beta_1 < 1$. Consequently, Theorem 4.4 is applicable.

5 First Example: A Phase Retrieval Problem

The problem of recovering a real-valued function, given only the amplitude but not the phase of its Fourier transform appears in applications to astronomy, electron microscopy, analysis of neutron reflectivity data and optical design (see [13], [18], [24], [31]). An introduction to the problem together with some descriptions of applications can be found in [29], [32]. For previous work on regularization methods for phase reconstruction we refer to [4], [5].

The phase retrieval problem can be formulated as operator equation (1) with the forward operator

$$\begin{aligned}
F : U \subseteq L^{\bar{p}}_{\mathbb{R}}(\mathbb{R}) &\to V = \tilde{V} = L^{\bar{p}_*}_{\mathbb{R}}(\mathbb{R}) \\
u &\mapsto |\mathcal{F}u| \,,
\end{aligned}$$

where $\mathcal{F}$ denotes the Fourier transform and

$$\bar{p} \in [1, 2] \text{ and } \frac{1}{\bar{p}} + \frac{1}{\bar{p}_*} = 1 \,.$$

Boundedness of F follows from well-known mapping properties of the Fourier transform, cf., e.g., [14]. Note that real valuedness of u implies a certain symmetry of its Fourier transform

$$u \in L^{\bar{p}}_{\mathbb{R}}(\mathbb{R}) \;\Rightarrow\; (\mathcal{F}u)(-s) = \overline{(\mathcal{F}u)(s)} \,.$$

The one-sided directional derivative of F is given by

$$F'(u^\dagger;h)(s) = \begin{cases} \dfrac{\Re((\mathcal{F}u^\dagger)(s)\overline{(\mathcal{F}h)(s))}}{|(\mathcal{F}u^\dagger)(s)|} \text{ if } s \in \mathbb{R}\setminus\Omega\,, \\ |(\mathcal{F}h)(s)| \text{ if } s \in \Omega\,, \end{cases} \tag{29}$$

where we define

$$\Omega = \{s \in \mathbb{R} : (\mathcal{F}u^\dagger)(s) = 0\}\,.$$

In both cases we have $|F'(u^\dagger;h)(s)| \le |(\mathcal{F}h)(s)|$, hence $F'(u^\dagger;h) \in L^{\bar{p}_*}$ follows from $\mathcal{F}h \in L^{\bar{p}_*}$. However, due to $\left|\mathcal{F}u^\dagger\right|$ appearing in the denominator, F cannot be expected to be Lipschitz continuously differentiable (as required in the literature on convergence rates for Tikhonov regularization so far). In this sense, we deal with a nonsmooth problem, as announced in the introduction.

We consider different sets U and regularization functionals depending on $\bar{p}$:

- If $\bar{p} \in (1,2]$ we take

$$\mathcal{D}(\mathcal{R}) = U = L^{\bar{p}}_{\mathbb{R}}(\mathbb{R})\,, \qquad \mathcal{R}(u) = \|u-u_0\|^2_{L^{\bar{p}}(\mathbb{R})}\,.$$

- For $\bar{p} = 1$ we use the negentropy regularization functional

$$\mathcal{R}(u) := \begin{cases} \displaystyle\int_{\mathbb{R}} \underbrace{(u-u_0)(\tau)\ln\left(\frac{u(\tau)-u_0(\tau)}{u_*(\tau)}\right)}_{=:g(\tau)} d\tau + \|u_*\|_{L^1(\mathbb{R})} \\ \quad \text{if } g \in L^1_{\mathbb{R}^+_0}(\mathbb{R}) \\ +\infty \qquad \text{else.} \end{cases} \tag{30}$$

$$U := L^1_{\mathbb{R}}(\mathbb{R})\,, \mathcal{D}(\mathcal{R}) = \{u \in L^1_{\mathbb{R}}(\mathbb{R}) \mid \mathcal{R} < \infty\}$$

  In here, u_0 is an initial guess and u_* an $L^1_{\mathbb{R}}(\mathbb{R})$ function with positive values.

  With the subgradient

$$\xi = 1 + \ln\left(\frac{u^\dagger - u_0}{u_*}\right)\,,$$

  the Bregman distance of $\mathcal{R}$ is given by

$$D_\xi(u,u^\dagger) = \int_{\mathbb{R}} \left\{(u(\tau)-u_0(\tau))\ln\left(\frac{u(\tau)-u_0(\tau)}{u^\dagger(\tau)-u_0(\tau)}\right) - (u(\tau)-u^\dagger(\tau))\right\} d\tau\,.$$

To be able to verify Assumption 2.1 in order to make use of the well–posedness, stability, and convergence results of Section 3, we use

$$\mathcal{D} := \text{ a sequentially compact subset of } U = L^{\bar{p}}_{\mathbb{R}}(\mathbb{R})$$

and use the strong topologies for defining τ_U, τ_V. For $\bar{p} \in (1,2]$ we use $V = L_{\mathbb{R}}^{\bar{p}_*}(\mathbb{R})$ and $\mathcal{R}(u) = \|u - u_0\|^2_{L^{\bar{p}}(\mathbb{R})}$; for $\bar{p} = 1$ we take $\mathcal{R}$ as in (30). In both cases Assumption 2.1 is satisfied.

As an example of sequentially compact sets in $L_{\mathbb{R}}^{\bar{p}}(\mathbb{R})$, $1 \leq \bar{p} < \infty$ we mention

$$\mathcal{D} := \left\{ u \in W_0^{1,1}([-a,a]) : \|u\|_{W_0^{1,1}([-a,a])} \leq C \right\}$$
$$\text{with } 0 < a < \infty \text{ and } C < \infty\,,$$

or in case $\bar{p} = 2$

$$\mathcal{D} := \left\{ u \in W^{\epsilon,2}([-a,a]) : \|u\|_{W^{\epsilon,2}([-a,a])} \leq C \right\}$$
$$\text{with } \epsilon > 0,\ 0 < a < \infty \text{ and } C < \infty.$$

In both cases we consider the functions to be extended by zero outside of $[-a,a]$. The first compactness result can be found in Adams [1, Theorem 6.2 on p144], the second one can be found in Lions & Magenes [35, Vol.1, Theorem 16.1. on p99]. Of course, under the compactness assumption well-posedness already follows from Tikhonov's Lemma even with $\alpha = 0$, but without convergence rates. Moreover, the analysis (without the rates) for the regularization methods with and without the addition of the penalization functional $\mathcal{R}$ is the same, since the regularization is already enforced by the compact set.

In view of this fact and Assumption 2.1, choosing the weak topology for defining τ_U and τ_V would suggest itself. (Actually, since L^1 is not reflexive, we would have to use the weak* topology in case $\bar{p} = 1$). However, a severe objection to the use of the weak topologies is that F is not continuous with respect to them. In $L^2(\mathbb{R})$, this can be seen by the simple counterexample

$$u_n(t) := \frac{1}{\sqrt{\pi}} \left(\frac{\sin(t - (2\pi n + \pi/2))}{t - (2\pi n + \pi/2)} + \frac{\sin(t + (2\pi n + \pi/2))}{t + (2\pi n + \pi/2)} \right)$$

whose Fourier transform is $(\mathcal{F}u_n)(s) = \sqrt{2}\cos((2\pi n + \pi/2)s)\chi_{[-1,1]}(s)$ which, as an orthonormal basis of $L^2([-1,1]$ weakly converges to $\mathcal{F}\tilde{u} \equiv 0$, hence $\tilde{u} \equiv 0$. However, with $w = \chi_{[-1,1]} \in L^2(\mathbb{R})$, we have

$$\int_{\mathbb{R}} F(u_n)(s) w(s)\, ds = \sqrt{2} \int_{-1}^{1} |\cos((2\pi n + \pi/2)s)|\, ds = \frac{\sqrt{2}}{\pi} \neq \int_{\mathbb{R}} F(\tilde{u})(s) w(s)\, ds\,.$$

Thus it seems that for the analysis, regularization by considering the solutions on a compact subset of $L_{\mathbb{R}}^{\bar{p}}(\mathbb{R})$ cannot be avoided.

In the sequel we verify the additional points in Assumptions 4.1, especially point 5.:

In order to formally derive the source condition (14) for this example, we rewrite on one hand

$$\langle \xi, h \rangle_{U^*,U} = \int_{\mathbb{R}} (\mathcal{F}\xi)(s)\overline{(\mathcal{F}h)(s)}\, ds = \int_{\mathbb{R}} \Re\Big((\mathcal{F}\xi)(s)\overline{(\mathcal{F}h)(s)} \Big)\, ds\,,$$

where we have applied Plancherel's theorem and the fact that the left hand side is real valued. On the other hand, we get

$$\begin{aligned}
&\left|\left\langle \tilde{\omega}, F'(u^\dagger; h)\right\rangle_{V^*,V}\right| \\
&= \left| \int_{\mathbb{R}\setminus\Omega} \tilde{\omega}(s) \frac{\Re((\mathcal{F}u^\dagger)(s)\overline{(\mathcal{F}h)(s)})\, ds}{|(\mathcal{F}u^\dagger)(s)|}\, ds + \int_\Omega \tilde{\omega}(s)\, |(\mathcal{F}h)(s)|\, ds \right| .
\end{aligned}$$

Therefore assuming that

$$\frac{\mathcal{F}\xi}{\mathcal{F}u^\dagger} \text{ is real valued on } \mathbb{R}\setminus\Omega \tag{31}$$

the source condition (17) (up to smallness $\beta_1 < 1$) is formally satisfied with

$$\tilde{\omega}(s) = \begin{cases} \left|(\mathcal{F}u^\dagger)(s)\right| \frac{(\mathcal{F}\xi)(s)}{(\mathcal{F}u^\dagger)(s)} & \text{if } s \in \mathbb{R}\setminus\Omega\,, \\ |(\mathcal{F}\xi)(s)| & \text{if } s \in \Omega\,. \end{cases} \tag{32}$$

Note that we here made use of the inequality option in this nonlinear version of a source condition, by estimating $\Re\Big((\mathcal{F}\xi)(s)\overline{(\mathcal{F}h)(s)}\Big) \leq |(\mathcal{F}\xi)(s)|\,|(\mathcal{F}h)(s)|$ for $s \in \Omega$. However, due to $\left|(\mathcal{F}u^\dagger)(s)\right|$ in the denominator of (29), condition (16) cannot be verified for this example. Therefore, we show (14) directly: For this purpose, we use the fact that

$$F(u^\dagger + h)(s) - F(u^\dagger)(s) = \begin{cases} \dfrac{2\Re((\mathcal{F}u^\dagger)(s)\overline{(\mathcal{F}h)(s)}) + |(\mathcal{F}h)(s)|^2}{|(\mathcal{F}(u^\dagger+h))(s)| + |(\mathcal{F}u^\dagger)(s)|} & \text{if } s \in \mathbb{R}\setminus\Omega\,, \\ |(\mathcal{F}h)(s)| & \text{if } s \in \Omega\,, \end{cases}$$

and rearrange as follows:

$$\begin{aligned}
&\left\langle \tilde{\omega}, F'(u^\dagger; h)\right\rangle_{V^*,V} \\
&= \int_{\mathbb{R}\setminus\Omega} \tilde{\omega}(s) \frac{\Re((\mathcal{F}u^\dagger)(s)\overline{(\mathcal{F}h)(s)})}{|(\mathcal{F}u^\dagger)(s)|}\, ds + \int_\Omega \tilde{\omega}(s)\, |(\mathcal{F}h)(s)|\, ds \\
&= \int_{\mathbb{R}\setminus\Omega} \frac{\tilde{\omega}(s)}{2\,|(\mathcal{F}u^\dagger)(s)|} \\
&\quad \left\{ \Big(F(u^\dagger+h) - F(u^\dagger)\Big)(s) \Big(\underbrace{\left|(\mathcal{F}(u^\dagger+h))(s)\right| + \left|(\mathcal{F}u^\dagger)(s)\right|}_{\leq 2|(\mathcal{F}u^\dagger)(s)| + |(\mathcal{F}h)(s)|} \Big) - |(\mathcal{F}h)(s)|^2 \right\} ds \\
&\quad + \int_\Omega \tilde{\omega}(s) \Big(F(u^\dagger+h) - F(u^\dagger)\Big)(s)\, ds\,,
\end{aligned}$$

hence by Hölder's inequality and $\|\mathcal{F}h\|_{L^{\bar{p}_*}(\mathbb{R})} \leq \|h\|_{L^{\bar{p}}(\mathbb{R})}$,

$$
\begin{aligned}
&\left|\left\langle \tilde{\omega}, F'(u^\dagger; h)\right\rangle_{V^*,V}\right| \\
\leq &\|\tilde{\omega}\|_{L^{\bar{p}}(\mathbb{R})} \left\|F(u^\dagger + h) - F(u^\dagger)\right\|_{L^{\bar{p}_*}(\mathbb{R})} \\
&+ \left\| \frac{\tilde{\omega}\,|\mathcal{F}h|}{2\,|\mathcal{F}u^\dagger|} \right\|_{L^{\bar{p}}(\mathbb{R}\setminus\Omega)} \left\|F(u^\dagger + h) - F(u^\dagger)\right\|_{L^{\bar{p}_*}(\mathbb{R})} \\
&+ \left\| \frac{\tilde{\omega}}{2\,|\mathcal{F}u^\dagger|} \right\|_{L^{\frac{1}{1-2/\bar{p}_*}}(\mathbb{R}\setminus\Omega)} \|\mathcal{F}h\|^2_{L^{\bar{p}_*}(\mathbb{R})} \\
\leq &\Big(\|\tilde{\omega}\|_{L^{\bar{p}}(\mathbb{R})} + \frac{1}{2} \left\| \frac{\tilde{\omega}}{|\mathcal{F}u^\dagger|} \right\|_{L^{\frac{\bar{p}}{1-\bar{p}/\bar{p}_*}}(\mathbb{R}\setminus\Omega)} \|h\|_{L^{\bar{p}}(\mathbb{R})} \Big) \left\|F(u^\dagger + h) - F(u^\dagger)\right\|_{L^{\bar{p}_*}(\mathbb{R})} \\
&+ \frac{1}{2} \left\| \frac{\tilde{\omega}}{|\mathcal{F}u^\dagger|} \right\|_{L^{\frac{1}{1-2/\bar{p}_*}}(\mathbb{R}\setminus\Omega)} \|h\|^2_{L^{\bar{p}}(\mathbb{R})} .
\end{aligned}
$$

Therewith, we can conclude that (14) holds with

$$
\begin{aligned}
\beta_1 =& \frac{1}{2} \bar{C}_1 \left\| \frac{\tilde{\omega}}{|\mathcal{F}u^\dagger|} \right\|_{L^{\frac{1}{1-2/\bar{p}_*}}(\mathbb{R}\setminus\Omega)} , \\
\beta_2 =& \|\tilde{\omega}\|_{L^{\bar{p}}(\mathbb{R})} + \frac{1}{2} \bar{C}_2 \left\| \frac{\tilde{\omega}}{|\mathcal{F}u^\dagger|} \right\|_{L^{\frac{\bar{p}}{1-\bar{p}/\bar{p}_*}}(\mathbb{R}\setminus\Omega)} ,
\end{aligned}
$$

provided we can establish estimates

$$
\forall u \in \mathcal{M}_{\alpha_{\max}}(\rho) \, : \, \left\|u - u^\dagger\right\|^2_{L^{\bar{p}}(\mathbb{R})} \leq \bar{C}_1 D_\xi(u, u^\dagger) \tag{33}
$$

$$
\forall u \in \mathcal{M}_{\alpha_{\max}}(\rho) \, : \, \left\|u - u^\dagger\right\|_{L^{\bar{p}}(\mathbb{R})} \leq \bar{C}_2 \tag{34}
$$

- In case $p \in (1, 2]$, (34) obviously holds with $\bar{C}_2 := \sqrt{\rho/\alpha_{\max}} + \left\|u_0 - u^\dagger\right\|$ and (33) follows from the proof of Corollary (ii) on page 192 in [10], (see also [11]).

  The derivative ξ used in the Bregman distance is given by

  $$
  \xi(s) = 2\left\|u^\dagger - u_0\right\|^{2-\bar{p}}_{L^{\bar{p}}(\mathbb{R})} (u^\dagger(s) - u_0(s))
  $$

  cf., e.g., Example 2.4 in [39].

- In case $p = 1$, estimate (34) directly follows from the fact that by the elementary estimate $\ln x \geq 1 - 1/x$, it follows that

$$
\begin{aligned}
\mathcal{R}(u) = &\begin{cases} \displaystyle\int_{\mathbb{R}} (u - u_0)(\tau) \left(1 - \frac{u_*(\tau)}{u(\tau) - u_0(\tau)}\right) d\tau + \int_{\mathbb{R}} u_*(\tau)\, d\tau \\ \qquad \text{if } g \in L^1_{\mathbb{R}_0^+}(\mathbb{R}) \\ +\infty \qquad \text{else} \end{cases} \\
\geq &\|u - u_0\|_{L^1(\mathbb{R})} .
\end{aligned} \tag{35}
$$

Relation (33) follows from results in [6], see also Proposition 2.12 in [12]: According to Lemma 2.2 in [6], we have

$$(y-x)^2 \;\le\; \left(\frac{2}{3}y+\frac{4}{3}x\right)\left(y\ln\frac{y}{x}-(y-x)\right)$$

for $x, y \in \mathbb{R}^+$, so setting $x(\tau)=\frac{u^\dagger(\tau)-u_0(\tau)}{u_*(\tau)}$, $y(\tau)=\frac{u(\tau)-u_0(\tau)}{u_*(\tau)}$, we get by the Cauchy-Schwarz inequality

$$\begin{aligned}
&\left\|u-u^\dagger\right\|_{L^1(\mathbb{R})}^2 = \left(\int_{\mathbb{R}} u_*(\tau)\left|y(\tau)-x(\tau)\right|\,d\tau\right)^2\\
&\le \left(\int_{\mathbb{R}}\sqrt{u_*(\tau)\left(\frac{2}{3}y(\tau)+\frac{4}{3}x(\tau)\right)}\sqrt{u_*(\tau)\left(y(\tau)\ln\frac{y(\tau)}{x(\tau)}-(y(\tau)-x(\tau))\right)}\,d\tau\right)^2\\
&\le \left(\frac{2}{3}\left\|u-u_0\right\|_{L^1(\mathbb{R})}+\frac{4}{3}\left\|u^\dagger-u_0\right\|_{L^1(\mathbb{R})}\right)D_\xi(u,u^\dagger)
\end{aligned}$$

Hence, by (35) and with $C_3(\lambda)=\frac{2}{3}\lambda+\frac{4}{3}\left\|u^\dagger-u_0\right\|_{L^1(\mathbb{R})}$, we get

$$\left\|u-u^\dagger\right\|_{L^1(\mathbb{R})}^2 \le C_3(\mathcal{R}(u))D_\xi(u,u^\dagger)\,,$$

which implies (33), due to the uniform bound $\mathcal{R}(u)\le\frac{\rho}{\alpha_{\max}}$ on elements u of $\mathcal{M}_{\alpha_{\max}}(\rho)$.
Recall that a derivative ξ of $\mathcal{R}$ can be defined by

$$\xi = 1+\ln\left(\frac{u^\dagger-u_0}{u_*}\right).$$

With the respective choice of the regularizing functional for $\bar{p}\in(1,2]$ and $\bar{p}=1$, respectively, $\tilde{\omega}$ given as in (32) satisfies condition (14), if (31) holds, and

$$\max\left\{\left\|\mathcal{F}(u^\dagger-u_0)\right\|_{L^{\bar{p}}(\mathbb{R})},\left\|\frac{\mathcal{F}(u^\dagger-u_0)}{|\mathcal{F}u^\dagger|}\right\|_{L^{\frac{\bar{p}}{1-\bar{p}/\bar{p}_*}}(\mathbb{R}\backslash\Omega)}\right\}<\infty$$

as well as

$$\left\|u^\dagger-u_0\right\|_{L^{\bar{p}}(\mathbb{R})}^{2-\bar{p}}\left\|\frac{\mathcal{F}(u^\dagger-u_0)}{|\mathcal{F}u^\dagger|}\right\|_{L^{\frac{1}{1-2/\bar{p}_*}}(\mathbb{R}\backslash\Omega)}\quad\text{sufficiently small}$$

in case $\bar{p}\in(1,2]$, and

$$\left\|\mathcal{F}(\ln(u^\dagger-u_0)-\ln(u_*/e))\right\|_{L^1(\mathbb{R})}<\infty$$

as well as

$$\left\|\frac{\mathcal{F}(\ln(u^\dagger-u_0)-\ln(u_*/e))}{|\mathcal{F}u^\dagger|}\right\|_{L^1(\mathbb{R}\backslash\Omega)}\quad\text{sufficiently small}$$

in case $\bar{p} = 1$, where we have used the identity

$$\xi = 1 + \ln\left(\frac{u^\dagger - u_0}{u_*}\right) = \ln(u^\dagger - u_0) - \ln(u_*/e)\,.$$

Note that in the latter case we end up with a closeness condition of $u^\dagger - u_0$ to $\frac{u_*}{e}$ rather than a closeness condition of $u^\dagger$ to u_0. Indeed, in case $\bar{p} = 1$ the purpose of u_0 is to ensure nonnegativity of $u - u_0$ but not necessarily to be a close approximation.

The real valuedness assumption (31) is indeed a quite strong one: In case $\bar{p} \in (1, 2]$, it implies that $\mathcal{F}u_0\overline{\mathcal{F}u^\dagger}$ is real valued and therewith, sloppily speaking, halves the dimension of the space of possible initial guesses u_0.

This assumption also had to be made in [4] to obtain convergence rates in a Hilbert space setting. Note however, that in order to be able to work in Hilbert spaces, we had to use stronger norms in [4] which resulted in considerably stronger smoothness conditions on $u^\dagger - u_0$ as compared to those assumed here.

6 Second Example: Incorporating Some Singular Case of Inverse Option Pricing

Inverse problems in option pricing and corresponding regularization approaches including convergence rates results have found increasing interest in the past ten years. Substantial contributions to that topic have been published by Bouchouev & Isakov [7, 8], Lagnado & Osher [34], Jackson, Süli & Howison [30], Crépey [17] and Egger & Engl [19] (see also [3, 16, 25, 36, 40]).

Therefore, as second example we reconsider a specific nonlinear inverse problem of this scenery, the problem of calibrating purely time-dependent volatility functions from maturity-dependent prices of European vanilla call options with fixed strike. We studied this problem in the papers [26] and [27]. It is certainly only a toy problem for mathematical finance, but due to its simple and explicit structure it serves as a benchmark problem for case studies in mathematical finance. However, following the decoupling approach suggested in [20] variants of this problem also occur as serious subproblems for the recovery of local volatility surfaces. Such surfaces are of considerable practical importance in finance.

For the inverse problem theory in Hilbert spaces the benchmark problem is of some interest, since in L^2 the standard convergence rates results of Tikhonov regularization for nonlinear ill-posed problems from [21, Chapter 10] cannot be applied for at-the-money options because of degeneration of the Fréchet derivative. This is just the case of options, where the degree of ill-posedness of the problem is minimal. We conjectured that the missing results in the case of frequently traded at-the-money options are only due to insufficient mathematical tools. For example in parameter identification problems of parabolic problems, where standard assumptions on Fréchet derivatives have not been available, Engl & Zou [23] could overcome this lack of smoothness of the forward operator by exploiting the inner structure of the problem. So we will also try to use the explicit character of our example problem. However, here we will work with auxiliary Banach spaces $L^{2-\varepsilon}$ in order to obtain rates results for the Hilbert space setting.

Precisely, our problem can be written as operator equation (1) in the Hilbert space

$$U = V = L^2(0,1).$$

At the present time we consider a family of European vanilla call options written on an asset with actual asset price $X > 0$ for varying maturities $t \in [0,1]$, but for a fixed strike price $K > 0$ and a fixed risk-free interest rate $r \geq 0$. We denote by $v(t)$ $(0 \leq t \leq 1)$ the associated function of option prices observed at an arbitrage-free financial market. From that function we are going to determine the unknown volatility term-structure. Furthermore, we denote the squares of the volatility at time t by $u(t)$ $(0 \leq t \leq 1)$ and neglect a possible dependence of the volatilities from asset price. Using a generalized Black-Scholes formula (see e.g. [33, p.71]) we obtain as the fair price function for the family of options

$$[F(u)](t) = U_{BS}(X, K, r, t, [J\,u](t)) \qquad (0 \leq t \leq 1) \tag{36}$$

using the simple integration operator

$$[J\,h]\,(s) \;=\; \int\limits_0^s h(t)\,dt \qquad (0 \leq s \leq 1)$$

and the Black-Scholes function U_{BS} defined as

$$U_{BS}(X, K, r, \tau, s) := \begin{cases} X\Phi(d_1) - Ke^{-r\tau}\Phi(d_2) & (s > 0) \\ \max\left(X - Ke^{-r\tau}, 0\right) & (s = 0) \end{cases}$$

with

$$d_1 = \frac{\ln\frac{X}{K} + r\tau + \frac{s}{2}}{\sqrt{s}}, \quad d_2 = d_1 - \sqrt{s}$$

and the cumulative density function of the standard normal distribution

$$\Phi(\zeta) \;=\; \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\zeta} e^{-\frac{\xi^2}{2}}\,d\xi.$$

We consider the nonlinear forward operator F of the problem mapping in $L^2(0,1)$, possessing the form (36) and having a natural convex domain

$$\mathcal{D}(F) \;=\; \left\{u \in L^2(0,1) :\; u(t) \geq c_0 > 0 \;\text{ a.e. on } [0,1]\right\}.$$

Obviously, the forward operator is a composition $F = N \circ J$ of the linear integral operator J with the nonlinear superposition operator

$$[N(z)](t) \;=\; k(t, z(t)) \;=\; U_{BS}(X, K, r, t, z(t)) \qquad (0 \leq t \leq 1)$$

having a smooth generator function k (cf. e.g. [2]).

Then with the linear multiplication operator

$$[G(u)\,h](t) = m(u,t)\,h(t) \qquad (0 \le t \le 1)$$

determined by a nonnegative multiplier function

$$m(u,0) = 0, \quad m(u,t) = \frac{\partial U_{BS}(X,K,r,t,[J\,u](t))}{\partial s} > 0 \qquad (0 < t \le 1)$$

(cf. [26, Lemma 2.1]), for which we can show for every $0 < t \le 1$ the formula

$$m(u,t) = \frac{X}{2\sqrt{2\pi[J\,u](t)}} \exp\left(-\frac{(\kappa + rt)^2}{2[J\,u](t)} - \frac{(\kappa + rt)}{2} - \frac{[J\,u](t)}{8}\right) > 0 \qquad (37)$$

with the logmoneyness $\kappa = \ln\left(\frac{X}{K}\right)$, the directional derivative is of the form $F'(u;h) = G(u)[J\,h]$ for $u \in \mathcal{D}(F)$ and $h \in L^2(0,1)$ and is characterized by the linear operator $G(u) \circ J$. So we can write for short $F'(u) = G(u) \circ J$. Note that in view of $c_0 > 0$ we have $\underline{c}\,t \le [J\,u](t) \le \overline{c}\,\sqrt{t}$ $(0 \le t \le 1)$ with $\underline{c} = c_0 > 0$ and $\overline{c} = \|u\|_{L^2(0,1)}$. Then we may estimate for all $u \in \mathcal{D}(F)$

$$\underline{C}\,\frac{\exp\left(-\frac{\kappa^2}{2\,\underline{c}\,t}\right)}{\sqrt[4]{t}} \;\le\; m(u,t) \;\le\; \overline{C}\,\frac{\exp\left(-\frac{\kappa^2}{2\,\overline{c}\,\sqrt{t}}\right)}{\sqrt{t}} \qquad (0 < t \le 1) \qquad (38)$$

with some positive constants $\underline{C}$ and $\overline{C}$.

If we exclude at-the money options, i.e. for

$$X \ne K \qquad (39)$$

and $\kappa := \ln\left(\frac{X}{K}\right) \ne 0$, the functions $m(u,\cdot)$ are continuous and have a uniquely determined zero at $t = 0$. In the neighborhood of this zero the multiplier function declines to zero exponentially, i.e. faster than any power of t, whenever the moneyness κ does not vanish (see formula (38)). From [26] we have in the case (39), where we either speak about in-the-money options or about out-of-the-money options, the following assertions: The multiplier functions $m(u,\cdot)$ all belong to $L^\infty(0,1)$ and hence $G(u)$ is a bounded multiplication operator in $L^2(0,1)$. Then $F'(u) = G(u) \circ J$ is a compact linear operator mapping in $L^2(0,1)$ and therefore a Gâteaux derivative for all $u \in \mathcal{D}(F)$. The nonlinear operator F is injective, continuous, compact, weakly continuous (and hence weakly closed) and $F'(u)$ is even a Fréchet derivative, for all $u \in \mathcal{D}(F)$, since it satisfies the condition

$$\|F(u_2) - F(u_1) - F'(u_1)\,(u_2 - u_1)\|_{L^2(0,1)} \;\le\; \gamma\,\|u_2 - u_1\|^2_{L^2(0,1)} \quad \text{for all } \; u_1, u_2 \in \mathcal{D}(F) \qquad (40)$$

with

$$\gamma = \frac{1}{2} \sup_{(t,s) \in [0,1]^2:\, 0 \le \underline{c}t \le s} \left|\frac{\partial^2 U_{BS}(X,K,r,t,s)}{\partial s^2}\right| < \infty\,. \qquad (41)$$

Note that γ, which can be interpreted as Lipschitz constant of $F'(u)$ for varying u, comes from the uniform boundedness of the second partial derivative of the Black-Scholes function U_{BS} with respect to the last variable, whereas the multiplier function $m(u,\cdot)$ defining $G(u)$ is due to the corresponding first partial derivative of U_{BS}.

As a consequence of the smoothing properties of F mentioned above the inverse operator F^{-1} : Range$(F) \subset L^2(0,1) \to L^2(0,1)$ exists, but cannot be continuous, and the corresponding operator equation (1) is locally ill-posed everywhere (in the sense of [28, Def. 2]). However, due to (40) the approach of [22] to analyzing the Tikhonov regularization with respect to convergence rates is directly applicable for the case $X \neq K$ and yields in that case the following proposition, which had been proven in [26, Theorem 5.4]. In the following we apply Theorem 4.4, where in our special situation $\mathcal{D}$ and $\mathcal{D}(F)$ coincide.

Proposition 6.1. *Provided that $X \neq K$ we have a convergence rate*

$$\|u_\alpha^\delta - u^\dagger\|_{L^2(0,1)} \;=\; O(\sqrt{\delta}) \qquad \text{as} \qquad \delta \to 0$$

for regularized solutions $u_\alpha^\delta \in \mathcal{D}$ of Tikhonov regularization minimizing the functional (2) with $p = 2$ and

$$\mathcal{R}(u) := \|u - u_0\|^2_{L^2(0,1)} \tag{42}$$

for some reference element $u_0 \in L^2(0,1)$ whenever the regularization parameter $\alpha = \alpha(\delta)$ is chosen a priori as $c\delta \le \alpha(\delta) \le C\delta$ $(0 < \delta \le \bar{\delta})$ for some positive constants c and C and the solution $u^\dagger$ of equation (1) fulfills the following two conditions, the first of which is a source condition and the second is a smallness condition. On the one hand, $u^\dagger$ has to satisfy

$$\xi = 2(u^\dagger - u_0) = F'(u^\dagger)^*\,\omega \tag{43}$$

for some $\omega \in L^2(0,1)$. The equation (43) implies that

$$u^\dagger(1) = u_0(1)\,. \tag{44}$$

If

$$(u^\dagger - u_0)' \text{ is measurable}$$

and the quotient function

$$\omega_0(t) := -\frac{2(u^\dagger - u_0)'(t)}{m(u^\dagger, t)} \quad (0 < t \le 1), \tag{45}$$

with $m(u,t)$ from (37), belongs to $L^2(0,1)$, then (43) holds true for $\omega = \omega_0$.

On the other hand, this function ω has to satisfy the inequality

$$\gamma\,\|\omega\|_{L^2(0,1)} \;<\; 1 \tag{46}$$

with γ from (41).

Moreover conditions (43), $\omega \in L^2(0,1)$ and the structure of $m(u^\dagger,\cdot)$ in the case $X \neq K$ imply that

$$u^\dagger - u_0 \in W^{1,2}(0,1)\,. \tag{47}$$

Proof. Since $1/m(u^\dagger, t)$ is measurable and we assume that $(u^\dagger - u_0)'$ is measurable, also ω_0 is measurable. It is easy to derive that the adjoint operator $F'(u^\dagger)^* : V^* = V \to U^* = U$ in (43) attains the form

$$F'(u^\dagger)^* = J^* \circ G(u^\dagger) : L^2(0,1) \to L^2(0,1).$$

Hence, the source condition (43) can be written as

$$2[u^\dagger - u_0](s) = \int_s^1 m(u^\dagger, t)\, \omega(t) dt \qquad (0 \le s \le 1)\,. \tag{48}$$

Then (44) follows from (48) by setting $s = 1$. By differentiation of (48) we obtain the quotient structure (45). Since $m(u,t)$ is a continuous function in $t \in [0,1]$ for all $u \in \mathcal{D}$ in the case $X \neq K$, based on (48) the property $\omega \in L^2(0,1)$ implies (47).

The sufficient conditions (17) and (16) of Remark 4.2, which are necessary to apply Theorem 4.4, follow immediately under the assumptions of this proposition taking into account Remark 4.3 and the estimate in (40). □

Taking into account the exponential order (see (38)) of the zero of $m(u^\dagger, t)$ at $t = 0$ it becomes evident that the source condition (43) and also the smallness condition (46) are strong requirements on the initial error $u^\dagger - u_0$ and its generalized derivative.

For at-the-money options with

$$X = K\,,$$

the proof of Proposition 6.1 along the lines of [26] unfortunately fails. We can also write $F'(u;h) = F'(u)\, h = G(u)[J\, h]$ for the directional derivative in that singular case $X = K$, and we have from [26, p.1322] (cf. also formula (37)) for the first partial derivative of the Black-Scholes function the explicit expression

$$\frac{\partial U_{BS}(X,K,r,t,s)}{\partial s} \;=\; \frac{X}{2\sqrt{2\,\pi\, s}} \exp\left(-\frac{r^2\, t^2}{2\, s} - \frac{r\, t}{2} - \frac{s}{8}\right) \,>\, 0 \tag{49}$$

and for the second partial derivative the expression

$$\begin{aligned} &\frac{\partial^2 U_{BS}(X,K,r,t,s)}{\partial s^2} = \\ &-\frac{X}{4\sqrt{2\,\pi\, s}} \left(-\frac{r^2\, t^2}{s^2} + \frac{1}{4} + \frac{1}{s}\right) \exp\left(-\frac{r^2\, t^2}{2\, s} - \frac{r\, t}{2} - \frac{s}{8}\right)\,. \end{aligned} \tag{50}$$

Moreover, in [27, p.55] based on formula (49) it was shown that the linear operator $F'(u) = G(u) \circ J$ is also bounded in the case $X = K$ and thus $F'(u)$ is even a Gâteaux derivative of F. However, by inspection of formula (50) we see that $\sup\limits_{(t,s)\in[0,1]^2:\, 0 \le \underline{c}t \le s} \left|\frac{\partial^2 U_{BS}(X,K,r,t,s)}{\partial s^2}\right| = \infty$. Hence, for vanishing moneyness $\kappa = 0$ an inequality (40) cannot be shown in such a way, since the constant γ

explodes. But (40) was required in [26] where Proposition 6.1 has been proven. On the other hand, for $\kappa = 0$ the forward operator F from (36) is less smoothing than for $\kappa \neq 0$, because for $X = K$ the multiplier function $m(u^\dagger, t)$ in $F'(u) = G(u) \circ J$ (also occurring in the denominator of (45)) has a pole at $t = 0$ for at-the-money options instead of a zero in all other cases. Hence, the local degree of ill-posedness of equation (1) in the singular case $X = K$ is smaller than in the regular case $X \neq K$. So it can be conjectured that an analogue of Proposition 6.1 also holds for $X = K$, but a more sophisticated approach for proving such a theorem is necessary and is given by Theorem 4.4 which allows us to compensate the degeneration of essential properties of the derivative $F'(u^\dagger)$.

In order to overcome the limitations of the singular situation with respect to convergence rates, we have to leave the pure Hilbert space setting. We directly apply Theorem 4.4 with $p = 2$, where we consider in addition to the Hilbert spaces $U = V = L^2(0,1)$ the Banach space

$$\tilde{V} = L^{2-\epsilon}(0,1) \supset V \qquad (0 < \epsilon < 1)$$

with dual space

$$\tilde{V}^* \text{isometrically isomorph to } L^{\frac{2-\epsilon}{1-\epsilon}}(0,1)\,,$$

again the stabilizing functional (42) defined on the whole space $L^2(0,1)$, which implies that

$$D_\xi(\tilde{u}, u) = \|\tilde{u} - u\|^2.$$

We assign the small value $\epsilon > 0$ a small value $\nu := \frac{\epsilon}{1-\epsilon} > 0$, where evidently we have $2 - \epsilon = \frac{2+\nu}{1+\nu}$ and $\frac{2-\epsilon}{1-\epsilon} = 2 + \nu$.

With the notation $S = J(u)$ and $S^\dagger = J(u^\dagger)$ we find for all $u \in \mathcal{D}$ the pointwise estimate

$$\begin{aligned}
&\left|[F(u) - F(u^\dagger) - F'(u^\dagger; u - u^\dagger)](t)\right| \\
&= \left|[F(u) - F(u^\dagger) - F'(u^\dagger)(u - u^\dagger)](t)\right| \\
&= \Bigg| U_{BS}(X,K,r,t,S(t)) - U_{BS}(X,K,r,t,S^\dagger(t)) - \\
&\qquad \frac{\partial U_{BS}(X,K,r,t,S^\dagger(t))}{\partial s}\left(S(t) - S^\dagger(t)\right)\Bigg| \\
&= \frac{1}{2}\left|\frac{\partial^2 U_{BS}(X,K,r,t,S_{im}(t))}{\partial s^2}\left(S(t) - S^\dagger(t)\right)^2\right|,
\end{aligned} \tag{51}$$

where S_{im} with $\min(S(t), S^\dagger(t)) \le S_{im}(t) \le \max(S(t), S^\dagger(t))$ is an intermediate function such that the pairs of real numbers $(t, S(t)), (t, S^\dagger(t))$ and $(t, S_{im}(t))$ all belong to the set $\{(t,s) \mid \underline{c}t \le s \; (0 \le t \le 1)\}$. Since

$$t\underline{c} \le S_{im}(t) \le t\|u_{im}\|_{L^2(0,1)} \quad (0 \le t \le 1)$$

it follows that we have for all t, s under consideration

$$\left| -\frac{X}{4\sqrt{2\,\pi\,s}} \quad \left(-\frac{r^2\,t^2}{s^2}+\frac{1}{4}+\frac{1}{s}\right) \exp\left(-\frac{r^2\,t^2}{2\,s}-\frac{r\,t}{2}-\frac{s}{8}\right)\right| \leq \\ \left| -\frac{X}{4\sqrt{2\,\pi\,s}} \quad \left(-\frac{r^2\,t^2}{s^2}+\frac{1}{4}+\frac{1}{s}\right)\right| \;\leq\; Ct^{-\frac{3}{2}}\;.$$

Thus we have shown that

$$\left| t\frac{\partial^2 U_{BS}(X,K,r,t,S_{im}(t))}{\partial s^2}\right| \leq Ct^{-\frac{1}{2}} \in L^{2-\epsilon}(0,1)\;.$$

Moreover, we have

$$\begin{aligned}\left(S(t)-S^\dagger(t)\right)^2 &= \left(\int_0^t (u(\tau)-u^\dagger(\tau)d\tau\right)^2 \\ &\leq \left(\int_0^t (u(\tau)-u^\dagger(\tau))^2 d\tau\right)\left(\int_0^t 1 d\tau\right)\end{aligned}$$

and thus

$$\begin{aligned}&\left\|F(u)-F(u^\dagger)-F'(u^\dagger;u-u^\dagger)\right\|^{2-\epsilon}_{L^{2-\epsilon}(0,1)} \\ =&\int_0^1 \Big| U_{BS}(X,K,r,t,S(t))-U_{BS}(X,K,r,t,S^\dagger(t))- \\ &\qquad \frac{\partial U_{BS}(X,K,r,t,S^\dagger(t))}{\partial s}\left(S(t)-S^\dagger(t)\right)\Big|^{2-\epsilon} dt \\ \leq &\int_0^1 \left|\frac{1}{2}\frac{\partial^2 U_{BS}(X,K,r,t,S_{im}(t))}{\partial s^2}\left(S(t)-S^\dagger(t)\right)^2\right|^{2-\epsilon} dt \\ \leq &\int_0^1 \left|\frac{1}{2}\frac{\partial^2 U_{BS}(X,K,r,t,S_{im}(t))}{\partial s^2}\right|^{2-\epsilon}\left|\frac{t}{2}\int_0^t (u(\tau)-u^\dagger(\tau))^2 d\tau\right|^{2-\epsilon} \\ \leq & C^{2-\epsilon}\left\| t\frac{\partial^2 U_{BS}(X,K,r,t,S_{im}(t))}{\partial s^2}\right\|^{2-\epsilon}_{L^{2-\epsilon}(0,1)}\left\|u-u^\dagger\right\|^{2(2-\epsilon)}_{L^2(0,1)}\;.\end{aligned}$$

This provides us with an estimate of the form (16) (here even valid for all $u \in \mathcal{D}$) with

$$\begin{aligned}\gamma &= C\left\| t\frac{\partial^2 U_{BS}(X,K,r,t,S_{im}(t))}{\partial s^2}\right\|_{L^{2-\epsilon}(0,1)} \\ &= C\left\| t\frac{\partial^2 U_{BS}(X,K,r,t,S_{im}(t))}{\partial s^2}\right\|_{L^{\frac{2+\nu}{1+\nu}}(0,1)}\end{aligned} \tag{52}$$

as required for Assumption 4.1. The estimate is the weaker analogue for the case $X = K$ of the estimate (40) which is valid for (39). Now we are ready to present the main theorem of this example.

Theorem 6.2. *We can extend the convergence rate assertion of Proposition 6.1 to the case $X = K$ if the assumptions of Proposition 6.1 hold true with the exception of the source condition and the smallness condition. For the source condition we assume in the case $X = K$ that*

$$(u^\dagger - u_0)' \text{ is measurable, } u^\dagger(1) = u_0(1),$$

and that there is an arbitrarily small $\nu > 0$ such that

$$\tilde{\omega}(t) := -\frac{2(u^\dagger - u_0)'(t)}{m(u^\dagger, t)} \quad (0 < t \le 1)$$

satisfies the condition

$$\tilde{\omega} \in L^{2+\nu}(0,1). \tag{53}$$

A smallness condition

$$\gamma \, \|\tilde{\omega}\|_{L^{2+\nu}(0,1)} < 1 \tag{54}$$

with γ from (52) has to be assumed.

Condition (53) *implies that*

$$u^\dagger - u_0 \in W^{1,1}(0,1).$$

Proof. Since $1/m(u^\dagger, t)$ is measurable and we assume that $(u^\dagger - u_0)'$ is measurable, also $\tilde{\omega}$ is measurable. We can estimate by Hölder's inequality and due to $\underline{c}\, t \le [Ju^\dagger](t) \le \overline{c}\, \sqrt{t}$ for $0 \le t \le 1$

$$\begin{aligned} \int_0^1 \left| m(u^\dagger, t)\, \tilde{\omega}(t) \right| dt &\le \left\| m(u^\dagger, t) \right\|_{L^{2-\epsilon}} \|\tilde{\omega}(t)\|_{L^{2+\nu}} \\ &\le \left\| C t^{-\frac{1}{2}} \right\|_{L^{2-\epsilon}} \|\tilde{\omega}(t)\|_{L^{2+\nu}} < \infty \end{aligned}$$

with some constant $C > 0$. Hence $2(u^\dagger - u_0)' = m(u^\dagger, \cdot)\, \tilde{\omega} \in L^1(0,1)$, and the function

$$[u^\dagger - u_0](s) = [u^\dagger - u_0](1) + \frac{1}{2} \int_s^1 m(u^\dagger, t) \tilde{\omega}(t) dt \qquad (0 \le s \le 1) \tag{55}$$

on the left-hand side of (55) is absolutely continuous and belongs to the Sobolev space $W^{1,1}(0,1)$.

Now for all $u \in \mathcal{D}$ and hence also for all $u \in \mathcal{M}_{\alpha_{max}}(\rho)$ we have for $\xi = 2(u^\dagger - u_0) \in \partial\mathcal{R}(u^\dagger)$

$$\begin{aligned}
\left\langle \xi, u - u^\dagger \right\rangle_{U^*,U} &= 2 \int\limits_0^1 \left(u^\dagger(t) - u_0(t)\right)\left(u(t) - u^\dagger(t)\right) dt \\
&= 2\left(u^\dagger(1) - u_0(1)\right) \int\limits_0^1 \left(u(t) - u^\dagger(t)\right) dt \\
&\quad - \int\limits_0^1 2\,(u^\dagger - u_0)'(t) \frac{m(u^\dagger, t)}{m(u^\dagger, t)} \left(\int_0^t \left(u(s) - u^\dagger(s)\right) ds \right) dt \\
&\le \left| \left\langle \tilde{\omega}, F'(u^\dagger; u - u^\dagger) \right\rangle_{V^*,V} \right| = \left| \left\langle \tilde{\omega}, F'(u^\dagger; u - u^\dagger) \right\rangle_{\tilde{V}^*,\tilde{V}} \right| .
\end{aligned}$$

The last equality holds since $F'(u^\dagger; u - u^\dagger) \in L^2(0,1)$.

For the proof of this theorem we again apply Theorem 4.4. The sufficient conditions (17) and (16) of Remark 4.2, which are to be shown in this context, follow immediately from the current assumptions, the equations in (51) and their consequences outlined above. □

Now it is an interesting task to compare the strength of source and smallness conditions in Theorem 6.2 and Theorem 6.1. After a rough inspection Theorem 6.2 seems to have stronger assumptions, because there is some additional $\nu > 0$. However, by a more precise inspection it gets clear that the requirements of (53), (54) for $X = K$ concerning the initial error $u^\dagger - u_0$ are much weaker than the requirements $\omega \in L^2(0,1)$ for ω from (45) and (46) on $u^\dagger - u_0$ in the case $X \neq K$. This is due to the exponential zero of $m(u^\dagger, t)$ at $t = 0$ for $X \neq K$ in contrast to a pole of m for $X = K$. More precisely, taking into consideration (49) we see that (53) and (54) hold if the function

$$(u^\dagger - u_0)'(t)\sqrt{[Ju^\dagger](t)}\exp\left(\frac{r^2\,t^2}{2\,[Ju^\dagger](t)} + \frac{r\,t}{2} + \frac{[Ju^\dagger](t)}{8}\right) \qquad (0 < t \le 1)$$

is in $L^{2+\nu}(0,1)$ and has there a sufficiently small norm. Then for the domain $\mathcal{D}$ under consideration the derivative of the initial error has to decay sufficiently fast near zero, namely as

$$(u^\dagger - u_0)'(t)\,\sqrt{[Ju^\dagger](t)} = O(t^{-\zeta}) \qquad \text{as} \qquad t \to 0$$

with $\zeta < \frac{1}{2+\nu}$. This condition is much weaker than the required exponential decay of the derivative of the initial error near $t = 0$ in the case $X \neq K$.

Finally, we can ask the question whether it is necessary to distinguish at all Proposition 6.1 in the regular case and Theorem 6.2 in the singular case, because Theorem 6.2 holds also true in the case $X \neq K$. However, it makes sense to formulate additionally the Proposition 6.1 for $X \neq K$, since $\nu > 0$ can be avoided there and the conditions (53) and (54) are stronger than the conditions $\omega \in L^2(0,1)$ and (46) which are appropriate for the regular case.

Acknowledgment

The second author thanks Elena Resmerita (RICAM, Linz) for her kind help in finding references [6], [10], [11] and [12] that contain crucial results for Section 5. Moreover, the authors thanks Heinz Engl (Linz) for pointing out several references related to the financial problems and the referees for their useful comments.

References

[1] R.A. Adams. *Sobolev Spaces.* Academic Press, New York, 1975.

[2] J. Appell and P.P. Zabrejko. *Nonlinear superposition operators.* Cambridge University Press, Cambridge, 1990.

[3] M. Avellaneda, C. Friedman, R. Holmes, and D. Samperi. Calibrating volatility surfaces via relative entropy minimization. *Appl. Math. Finance*, 4:3–64, 1997.

[4] B. Blaschke, H.W. Engl, W. Grever, and M.V. Klibanov. An application of Tikhonov regularization to phase retrieval. *Nonlinear World*, 3:771–786, 1996.

[5] B. Blaschke-Kaltenbacher and H.W. Engl. Regularization methods for nonlinear ill–posed problems with applications to phase reconstruction. In H.W. Engl, A.K. Louis, and W.Rundell, editors, *Inverse Problems in Medical Imaging and Nondestructive Testing*, pages 1–18, Wien-New York, 1997. Springer.

[6] J. M. Borwein and A. S. Lewis. Convergence of best entropy estimates. *SIAM J. Optim.*, 1:191–205, 1991.

[7] I. Bouchouev and V. Isakov. The inverse problem of option pricing. *Inverse Problems*, 13:L1–L7, 1997.

[8] I. Bouchouev and V. Isakov. Uniqueness, stability and numerical methods for the inverse problem that arises in financial markets. *Invere Problems*, 15:R95–116, 1999.

[9] M. Burger and S. Osher. Convergence rates of convex variational regularization. *Inverse Problems*, 20:1411–1421, 2004.

[10] D. Butnariu and A.N. Iusem. *Totally Convex Functions for Fixed Points Computation and Infinite Dimensional Optimization.* Kluwer, Dordrecht, 2000.

[11] D. Butnariu and E. Resmerita. The outer Bregman projection method for solving stochastic convex feasibility problems in Banach spaces. In D. Butnariu, Y. Censor, and S. Reich, editors, *Inherently Parallel Algorithms in*

Feasibility and Optimization and their Applications, pages 69–86, Amsterdam, 2001. Elsevier.

[12] D. Butnariu and E. Resmerita. Bregman distances, totally convex functions and a method for solving operator equations. *Abstract and Applied Analysis*, 2006.

[13] K. Chadan and P.C. Sabatier. *Inverse Problems in Quantum Scattering Theory*. Springer, New York, 1989.

[14] D.C. Champeney. *A Handbook of Fourier Theorems*. Cambridge University Press, Cambridge, 1987.

[15] G. Chavent and K. Kunisch. Regularization of linear least squares problems by total bounded variation. *ESAIM Control Optim. Calc. Var.*, 2:359–376 (electronic), 1997.

[16] R. Cont and P. Tankov. Retrieving Lévy processes from option prices: Regularization of an il-posed inverse problem. *SIAM J. Control Optim.*, 45:1–25, 2006.

[17] S. Crépey. Calibration of the local volatility in a generalized Black-Scholes model using Tikhonov regularization. *SIAM J. Math. Anal.*, 34:1183–1206, 2003.

[18] D. Dobson. Phase reconstruction via nonlinear least squares. *Inverse Problems*, 8:541–557, 1992.

[19] H. Egger and H. W. Engl. Tikhonov regularization applied to the inverse problem of option pricing: convergence analysis and rates. *Inverse Problems*, 21:1027–1045, 2005.

[20] H. Egger, T. Hein, and B. Hofmann. On decoupling of volatility smile and term structure in inverse option pricing. *Inverse Problems*, 22:1247–1259, 2006.

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

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

[23] H.W. Engl and J. Zou. A new approach to convergence rate analysis of Tikhonov regularization. *Inverse Problems*, 16:1907–1923, 2000.

[24] G. Felcher and T. Russell. *Proceedings of the Workshop on Methods of Analysis and Interpretation of Neutron Reflectivity Data*. Phys. B., Vol. 173, 1991.

[25] M. Hanke and E. Rösler. Computation of local volatilities from regularized Dupire equations. *Int. J. Theor. Appl. Finance*, 8:207–222, 2005.

[26] T. Hein and B. Hofmann. On the nature of ill-posedness of an inverse problem arising in option pricing. *Inverse Problems*, 19:1319–1338, 2003.

[27] B. Hofmann and R. Krämer. On maximum entropy regularization for a specific inverse problem of option pricing. *J. Inv. Ill-Posed Problems*, 13:41–63, 2005.

[28] B. Hofmann and O. Scherzer. Factors influencing the ill-posedness of nonlinear problems. *Inverse Problems*, 10:1277–1297, 1994.

[29] N.E. Hurt. *Phase Retrieval and Zero Crossings.* Kluwer, Dordrecht, 1989.

[30] N. Jackson, E. Süli, and S. Howison. Computation of deterministic volatility surfaces. *J. Mathematical Finance*, 2:5–32, 1998.

[31] M.V. Klibanov and P.E. Sacks. Phaseless inverse scattering and the phase problem in optics. *J. Math. Phys.*, 33:3913–3821, 1992.

[32] M.V. Klibanov, P.E. Sacks, and A.V. Tikhonravov. The phase retrieval problem. *Inverse Problems*, 11:1–28, 1995.

[33] Y. K. Kwok. *Mathematical Models of Financial Derivatives.* Springer, Singapore, 1998.

[34] R. Lagnado and S. Osher. A technique for calibrating derivative security pricing models: numerical solution of the inverse problem. *J. Computational Finance*, 1:13–25, 1997.

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

[36] J. Lishang and T. Youshan. Identifying the volatility of underlying assets from option prices. *Inverse Problems*, 17:137–155, 2001.

[37] V.A. Morozov. *Methods for Solving Incorrectly Posed Problems.* Springer Verlag, New York, Berlin, Heidelberg, 1984.

[38] E. Resmerita and O. Scherzer. Error estimates for non-quadratic regularization and the relation to enhancing. *Inverse Problems*, 22:801–814, 2006.

[39] F. Schöpfer, A.K. Louis, and T. Schuster. Nonlinear iterative methods for linear ill-posed problems in Banach space. *Inverse Problems*, 22:311–329, 2006.

[40] J. P. Zubelli. Inverse problems in finance: A short survey of calibration techniques. In N. Kolev and P. Morettin, editors, *Proc. 2nd Brazilian Conference on Statistical Modelling in Insurance and Finance, Maresias, Brazil, 2005*, pages 64–75. Institute of Mathematics and Statistics, University of São Paulo, 2005.