

Existence of variational source conditions for nonlinear inverse problems in Banach spaces

JENS FLEMMING*

September 21, 2017

Key words: ill-posed problem, convergence rates, variational source condition, nonlinear equation, Banach space

MSC: 65J20, 47J06

Abstract

Variational source conditions proved to be useful for deriving convergence rates for Tikhonov's regularization method and also for other methods. Up to now such conditions have been verified only for few examples or for situations which can be handled by classical range-type source conditions, too.

Here we show that for almost every ill-posed inverse problem variational source conditions are satisfied. Whether linear or nonlinear, whether Hilbert or Banach spaces, whether one or multiple solutions, variational source conditions are a universal tool for proving convergence rates.

1 Setting

We want to solve nonlinear equations

$$F(x) = y^\dagger, \quad x \in X, \quad (1.1)$$

in Banach spaces X and Y , where $F : D(F) \rightarrow Y$ has domain $D(F) \subseteq X$. The exact right-hand side y^\dagger may be known only approximately as a noisy measurement $y^\delta \in Y$ with

$$\|y^\delta - y^\dagger\| \leq \delta$$

*Chemnitz University of Technology, Faculty of Mathematics, D-09107 Chemnitz, Germany, jens.flemming@mathematik.tu-chemnitz.de.

for some noise level $\delta \geq 0$. In addition, the solutions need not depend continuously on this data.

Regularization is required and we concentrate on Tikhonov's method

$$\|F(x) - y^\delta\|^p + \alpha \Omega(x) \rightarrow \min_{x \in D(F)}. \quad (1.2)$$

We assume $p \geq 1$ and $\alpha > 0$. The penalty $\Omega : X \rightarrow (-\infty, \infty]$ is allowed to attain $+\infty$ as a value. This method is well understood and existence, stability and convergence of minimizers can be guaranteed by the following assumptions, see [22, Section 3.2] or [23, Chapter 4].

Assumption 1. We assume that the following properties are satisfied by the introduced setting.

- (i) Equation (1.1) has a solution with finite Ω .
- (ii) F is weakly sequentially closed, that is, each sequence $(x_n)_{n \in \mathbb{N}}$ in $D(F)$ satisfies

$$x_n \rightharpoonup x, \quad F(x_n) \rightharpoonup y \quad \Rightarrow \quad x \in D(F), \quad F(x) = y.$$

- (iii) Ω is proper and convex.
- (iv) The sublevel sets $\{x \in X : \Omega(x) \leq c\}$, $c \in \mathbb{R}$, are weakly sequentially closed and each sequence in such a set has a weakly convergent subsequence.

Tikhonov regularized solutions always converge weakly, at least in a subsequential manner, to solutions which minimize the penalty Ω in the set of all solutions. Ω -minimizing solutions are typically denoted by x^\dagger . Convergence to such Ω -minimizing solutions may be arbitrarily slow and we are interested in estimates for the convergence speed. Of course, additional assumptions are required for such estimates. This issue will be discussed in the next section.

At first, we have to decide how to measure the speed of convergence. To cover all relevant situations we introduce a general error functional $E^\dagger : X \rightarrow [0, \infty)$, where the symbol \dagger indicates that the error functional depends on one or more Ω -minimizing solutions. The canonical choice in Hilbert spaces is the squared norm distance

$$E^\dagger(x) = \|x - x^\dagger\|^2$$

between regularized and exact solution. If there are multiple solutions, the point-to-set distance

$$E^\dagger(x) = \inf_{x^\dagger \in S} \|x - x^\dagger\|^2$$

is a suitable choice if S denotes the set of norm minimizing solutions. In Banach spaces alternatives like the Bregman distance

$$E^\dagger(x) = B_{\xi^\dagger}(x, x^\dagger)$$

(see Section 4 for a definition) proved to be useful. But Banach space norms can also be used. In $\ell^1(\mathbb{N})$ we could choose

$$E^\dagger(x) = \|x - x^\dagger\|_1 \quad \text{or} \quad E^\dagger(x) = \inf_{x^\dagger \in S} \|x - x^\dagger\|_1$$

with S denoting again the set of norm minimizing solutions. All these examples will be discussed in Section 4 in more detail.

Denoting by x_α^δ the minimizers of the Tikhonov minimization problem (1.2), we aim at asymptotic estimates

$$E^\dagger(x_\alpha^\delta) = \mathcal{O}(\varphi(\delta)), \quad \delta \rightarrow 0, \quad (1.3)$$

where α may depend on δ and y^δ . The function φ shall be an index function in the following sense.

Definition 2. A function $\varphi : [0, \infty) \rightarrow [0, \infty)$ is an *index function* if it is continuous, monotonically increasing, strictly increasing in a neighborhood of zero, and satisfies $\varphi(0) = 0$.

2 Variational source conditions revisited

Different techniques have been developed to prove convergence rates (1.3). The most prominent tool are source conditions for linear ill-posed inverse problems in Hilbert spaces. The classical concept is described in [6, Section 3.2] and general source conditions are studied in [21]. See also the references given in [21] for the origins of general source conditions. In both cases the norm distance between exact and regularized solution is used as error functional E^\dagger .

For Banach spaces usage of source conditions is quite limited. But in 2007 variational source conditions were introduced in [14] and thoroughly studied and developed during the past 10 years, see, e.g., [2, 7, 11, 17]. This type of condition allows to prove convergence rates for many different

settings, especially for nonlinear operators and general penalty functionals in (1.2). In its original version Bregman distances (see Section 4) were used as error functional E^\dagger .

Variational source conditions are also known as variational inequalities, but this term conflicts with the already existing mathematical field with the same name. A second alternative was introduced in the book [7]. There the term variational smoothness assumption is used, because several kinds of smoothness (not only of the underlying exact solution as it is the case for classical source conditions) are jointly described by one expression. The term variational source condition rouses associations to classical source conditions. But the new concept has no similarity to classical source conditions, most notably there is no source element. Nevertheless, in most recent literature ‘variational source condition’ seems to be used more often than ‘variational inequality’, whereas ‘variational smoothness assumption’ is not used by other authors. Thus, we write ‘variational source condition’ to name the technique described and applied below and, to avoid drawing too many parallels to source conditions, we read it as ‘variational replacement for source conditions’.

The definition of variational source conditions in the present paper will be slightly more general than other variants before, because it is not connected to one fixed solution of (1.1). Instead, we allow multiple Ω -minimizing solutions and refer to [5] for a concrete example. For this purpose we denote by Ω^\dagger the value of Ω at the Ω -minimizing solutions, that is,

$$\Omega^\dagger := \min\{\Omega(x^\dagger) : x^\dagger \in D(F), F(x^\dagger) = y^\dagger\}.$$

Assumption 1 guarantees existence of Ω -minimizing solutions.

Definition 3. Let $\beta > 0$ be a constant and let $\varphi : [0, \infty) \rightarrow [0, \infty)$ be an index function. A *variational source condition* for fixed right-hand side y^\dagger holds on a set $M \subseteq D(F)$ if

$$\beta E^\dagger(x) \leq \Omega(x) - \Omega^\dagger + \varphi(\|F(x) - y^\dagger\|) \quad \text{for all } x \in M. \quad (2.1)$$

If the set M is large enough to contain all minimizers of the Tikhonov functional (1.2), then a variational source condition (2.1) implies the desired convergence rate (1.3). Although our variant is slightly more general, the proofs of this fact given in [7, Chapter 4] or in [15] still work with trivial modifications. Suitable choices of the regularization parameter α are discussed there, too.

The constant β plays only a minor role. In principle we could hide it in the functional E^\dagger , but then E^\dagger would depend on the chosen index function φ

and not solely on exact and regularized solutions. The implied convergence rate does not depend on β , only the \mathcal{O} -constant contains the factor $\frac{1}{\beta}$.

Variational source conditions originally were developed to obtain rates for Tikhonov regularization, but can also be used in the context of other methods. See [13] for the residual method and [17] for iteratively regularized Newton methods.

A major drawback of variational source conditions is that the best obtainable rate may be slower than the best possible one. This is for instance the case for rates faster than $\mathcal{O}(\sqrt{\delta})$ in the classical linear Hilbert space setting, where the best one is $\mathcal{O}(\delta^{\frac{2}{3}})$. On the other hand, in ℓ^1 -regularization rates up to the best possible one $\mathcal{O}(\delta)$ for the error norm can be obtained, see [3]. An approach to overcome technical rate limitations was undertaken in [12], but it is limited to linear equations.

3 Main result

The aim of this paper is to answer the question under which conditions we find a variational source condition of type (2.1). For linear equations in Hilbert spaces we know that general source conditions can always be satisfied, see [20], and that general source conditions imply variational source conditions without reducing the implied convergence rate, see [7, Chapter 13]. In Banach spaces up to now variational source conditions were only verified for few concrete examples, see, e. g., [5, 14, 19]. The only exception is ℓ^1 -regularization for linear equations, where in [9] it was shown that one always finds a variational source condition for some index function φ .

Of course, we have to connect the error functional E^\dagger in some way to the other ingredients of a variational source condition.

Assumption 4. Given $\beta > 0$ and $M \subseteq D(F)$ we assume the following.

- (i) M is weakly sequentially closed.
- (ii) All solutions x^* to (1.1) satisfy the inequality in (2.1) if they belong to M , that is,

$$\beta E^\dagger(x^*) \leq \Omega(x^*) - \Omega^\dagger$$

for all $x^* \in M$ with $F(x^*) = y^\dagger$.

- (iii) The mapping

$$x \mapsto -\beta E^\dagger(x) + \Omega(x)$$

is weakly sequentially lower semicontinuous on M .

(iv) There are constants $\tilde{\beta} > \beta$ and $\tilde{c} \geq 0$ such that

$$\tilde{\beta} E^\dagger(x) - \Omega(x) \leq \tilde{c} \quad \text{for all } x \in M.$$

These assumptions will be verified for several important special cases in Section 4. There we will see that all relevant settings are covered.

Theorem 5. *Let Assumption 1 and Assumption 4 be true for a constant $\beta > 0$ and a set $M \subseteq D(F)$. Then there exists a concave index function φ such that the variational source condition (2.1) is satisfied.*

The assertion of the theorem is quite similar to the main result of [20], where it is shown that in linear Hilbert space settings one always finds an index function such that a corresponding general source condition is satisfied. Our theorem extends this results to nonlinear Banach space settings.

Remark 6. As already noted in Section 2, a variational source condition (2.1) only implies the desired convergence rate (1.3) if the set M is large enough. Here, ‘large enough’ means that all regularized solutions have to belong to M , at least for small noise levels δ . In Theorem 5 the size of M is restricted only by Assumption 4 and, as we will see in Section 4, Assumption 4 is usually satisfied with $M = D(F)$. Thus, Theorem 5 not only yields some artificial variational source condition, but a variational source condition which implies rates.

The proof of the theorem relies on the technique of approximate variational source conditions introduced in [10] and thoroughly studied in [7, Chapter 12]. Here we only introduce the parts we need for the proof. The idea is to measure the violation of a variational source condition with linear index function φ by the *distance function*

$$D_\beta(r) := \sup_{x \in M} \left(\beta E^\dagger(x) - \Omega(x) + \Omega^\dagger - r \|F(x) - y^\dagger\| \right) \quad (3.1)$$

for $r \geq 0$. This function is the supremum of affine functions and thus convex and continuous on the interior of its domain. Obviously, it is monotonically decreasing.

Lemma 7. *Let D_β be defined by (3.1). If $D_\beta(0) > 0$ and*

$$\lim_{r \rightarrow \infty} D_\beta(r) = 0,$$

then

$$\varphi(t) = \inf_{r \geq 0} (D_\beta(r) + r t), \quad t \geq 0,$$

defines a concave index function and a variational source condition (2.1) with this φ and same β and M as in (3.1) holds true.

Proof. From

$$\begin{aligned} & \beta E^\dagger(x) - \Omega(x) + \Omega^\dagger \\ &= \inf_{r \geq 0} (\beta E^\dagger(x) - \Omega(x) + \Omega^\dagger - r \|F(x) - y^\dagger\| + r \|F(x) - y^\dagger\|) \\ &\leq \inf_{r \geq 0} (D_\beta(r) + r \|F(x) - y^\dagger\|) \end{aligned}$$

for all $x \in M$ we obtain the asserted variational source condition if φ is indeed an index function.

Since D_β is decreasing and goes to zero, it has to be nonnegative. Consequently, $0 \leq \varphi(t) < \infty$. We also immediately see that φ is monotonically increasing. In addition, φ is concave and upper semicontinuous as an infimum of affine functions. Thus, φ is continuous on the interior $(0, \infty)$ of its domain. The decay of D_β to zero yields $\varphi(0) = 0$, which, together with upper semicontinuity, yields continuity on the whole domain $[0, \infty)$. The assumption $D_\beta(0) > 0$ ensures $\varphi(t) > 0$ for some t . Thus, φ has to be strictly increasing near zero. \square

Proof of Theorem 5. We want to apply Lemma 7. If $D_\beta(0) \leq 0$, then the variational source condition holds for arbitrary index functions φ . So we may assume $D_\beta(0) > 0$ and the lemma reduces the proof to verification of $D_\beta(r) \rightarrow 0$ if $r \rightarrow \infty$. In addition, we may assume $D_\beta(r) \geq 0$ for all $r > 0$, because $D_\beta(r) < 0$ for some r would imply a variational source condition with the best possible concave index function $\varphi(t) = rt$.

As first step we show that for fixed $r \geq 0$ the supremum in the definition (3.1) of $D_\beta(r)$ is attained. We write

$$D_\beta(r) := - \inf_{x \in M} \left(-\beta E^\dagger(x) + \Omega(x) - \Omega^\dagger + r \|F(x) - y^\dagger\| \right) \quad (3.2)$$

and denote by $(x_n)_{n \in \mathbb{N}}$ an infimizing sequence. The functional in the infimum is bounded on the sequence by some constant c . With $\tilde{\beta}$ and \tilde{c} as in Assumption 4 (iv) we see

$$\begin{aligned} -\beta E^\dagger(x_n) + \Omega(x_n) &= \frac{\beta}{\tilde{\beta}} (-\tilde{\beta} E^\dagger(x_n) + \Omega(x_n)) + \left(1 - \frac{\beta}{\tilde{\beta}}\right) \Omega(x_n) \\ &\geq \frac{\beta \tilde{c}}{\tilde{\beta}} + \left(1 - \frac{\beta}{\tilde{\beta}}\right) \Omega(x_n), \end{aligned}$$

which implies

$$\begin{aligned}
\left(1 - \frac{\beta}{\tilde{\beta}}\right) \Omega(x_n) &\leq -\beta E^\dagger(x_n) + \Omega(x_n) - \frac{\beta \tilde{c}}{\tilde{\beta}} \\
&\leq -\beta E^\dagger(x_n) + \Omega(x_n) - \Omega^\dagger + r \|F(x_n) - y^\dagger\| + \Omega^\dagger - \frac{\beta \tilde{c}}{\tilde{\beta}} \\
&\leq c + \Omega^\dagger - \frac{\beta \tilde{c}}{\tilde{\beta}}.
\end{aligned}$$

Thus, $(\Omega(x_n))_{n \in \mathbb{N}}$ is bounded and we find a subsequence of $(x_n)_{n \in \mathbb{N}}$ converging weakly to some $\tilde{x} \in X$. The subsequence will be denoted again by $(x_n)_{n \in \mathbb{N}}$. Assumption 4 (i) ensures $\tilde{x} \in M$ and with item (iii) one easily shows that \tilde{x} is a minimizer in (3.2).

As second step we take a sequence $(r_n)_{n \in \mathbb{N}}$ in $[0, \infty)$ with $r_n \rightarrow \infty$ and corresponding maximizers x_n in the definition (3.1) of $D_\beta(r_n)$ to show $D_\beta(r) \rightarrow \infty$ if $r \rightarrow \infty$. As discussed above, Assumption 4 (iv) implies

$$\begin{aligned}
\left(1 - \frac{\beta}{\tilde{\beta}}\right) \Omega(x_n) &\leq -\beta E^\dagger(x_n) + \Omega(x_n) - \Omega^\dagger + r_n \|F(x_n) - y^\dagger\| + \Omega^\dagger - \frac{\beta \tilde{c}}{\tilde{\beta}} \\
&= -D_\beta(r_n) + \Omega^\dagger - \frac{\beta \tilde{c}}{\tilde{\beta}} \\
&\leq \Omega^\dagger - \frac{\beta \tilde{c}}{\tilde{\beta}}.
\end{aligned}$$

Thus, $(\Omega(x_n))_{n \in \mathbb{N}}$ is bounded and we find a subsequence of $(x_n)_{n \in \mathbb{N}}$ converging weakly to some $\tilde{x} \in M$. The subsequence will be denoted again by $(x_n)_{n \in \mathbb{N}}$ with corresponding $(r_n)_{n \in \mathbb{N}}$.

From $D_\beta(r_n) \geq 0$ we obtain

$$r_n \|F(x_n) - y^\dagger\| \leq \beta E^\dagger(x_n) - \Omega(x_n) + \Omega^\dagger$$

and the right-hand side is bounded by Assumption 4 (iv). Consequently, $r_n \rightarrow \infty$ implies $\|F(x_n) - y^\dagger\| \rightarrow 0$ and the lower semicontinuity of $x \mapsto \|F(x) - y^\dagger\|$ yields $F(\tilde{x}) = y^\dagger$, that is, the maximizers in the definition of D_β converge (subsequentially) to solutions of (1.1). If we combine this

observation with items (ii) and (iii) in Assumption 4, we obtain

$$\begin{aligned}
0 &\leq \liminf_{n \rightarrow \infty} D_\beta(r_n) \leq \limsup_{n \rightarrow \infty} D_\beta(r_n) = -\liminf_{n \rightarrow \infty} (-D_\beta(r_n)) \\
&= -\liminf_{n \rightarrow \infty} (-\beta E^\dagger(x_n) + \Omega(x_n) - \Omega^\dagger + r_n \|F(x_n) - y^\dagger\|) \\
&\leq -\liminf_{n \rightarrow \infty} (-\beta E^\dagger(x_n) + \Omega(x_n) - \Omega^\dagger) \\
&\leq -(-\beta E^\dagger(\tilde{x}) + \Omega(\tilde{x}) - \Omega^\dagger) \\
&\leq 0.
\end{aligned}$$

This proves $D_\beta(r_n) \rightarrow 0$ and, since $(r_n)_{n \in \mathbb{N}}$ was chosen arbitrarily, also $D_\beta(r) \rightarrow 0$ if $r \rightarrow \infty$. The fact that $(r_n)_{n \in \mathbb{N}}$ is only a subsequence of the original sequence causes no troubles, because D_β is monotonically decreasing. Thus, $D_\beta(r_n) \rightarrow 0$ has to hold for the original sequence, too. \square

Theorem 5 states that there is always an index function φ for a variational source condition. In principle the proof is constructive, but calculating the distance function D_β is a difficult task. From the proof of Lemma 7 we see that also a majorant for D_β yields an index function φ as long as this majorant decays to zero at infinity. Index functions φ constructed in this spirit can be found in [5, 9] for special settings. A very similar approach to obtain index functions φ is discussed in [16].

4 Special cases

We provide some special cases for which Theorem 5 is applicable.

4.1 Linear equations in Hilbert spaces

Let X and Y be Hilbert spaces and let F be linear and bounded with $D(F) = X$. For the Tikhonov functional (1.2) we choose $p = 2$ and $\Omega(x) = \|x\|^2$. For each α there is exactly one Tikhonov minimizer. Equation (1.1) may have multiple solutions, but there is exactly one Ω -minimizing solution, which we denote by x^\dagger . As error functional E^\dagger we choose

$$E^\dagger(x) = \|x - x^\dagger\|^2, \quad x \in X.$$

Corollary 8. *For each $\beta \in (0, 1)$ there exists a concave index function φ such that the variational source condition (2.1) is satisfied with $M = X$.*

Proof. Assumption 1 is obviously true and we only have to check Assumption 4 to apply Theorem 5. Item (i) is trivially true. Item (ii) reads

$$\beta \|x^* - x^\dagger\|^2 \leq \|x^*\|^2 - \|x^\dagger\|^2 \quad \text{for all } x^* \in X \text{ with } Ax^* = y^\dagger.$$

For solutions x^* the difference $x^* - x^\dagger$ is in the nullspace of A and x^\dagger is in the orthogonal complement of the null space. Thus, the inner product $\langle x^\dagger, x^* - x^\dagger \rangle$ vanishes and we obtain the desired estimate

$$\begin{aligned} \beta \|x^* - x^\dagger\|^2 &= \beta (\|x^*\|^2 - \|x^\dagger\|^2 - 2 \operatorname{Re} \langle x^\dagger, x^* - x^\dagger \rangle) \\ &= \beta (\|x^*\|^2 - \|x^\dagger\|^2) \\ &\leq \|x^*\|^2 - \|x^\dagger\|^2. \end{aligned}$$

Concerning item (iii) we observe

$$-\beta \|x - x^\dagger\|^2 + \|x\|^2 = (1 - \beta) \|x\|^2 - 2\beta \operatorname{Re} \langle x, x^\dagger \rangle + \beta \|x^\dagger\|^2$$

for all $x \in X$. Since $\beta < 1$ by assumption, the functional is obviously weakly sequentially lower semicontinuous. Finally, item (iv) follows from

$$\begin{aligned} \tilde{\beta} \|x - x^\dagger\| - \|x\|^2 &= -(1 - \tilde{\beta}) \|x\|^2 - 2\tilde{\beta} \operatorname{Re} \langle x, x^\dagger \rangle + \tilde{\beta} \|x^\dagger\|^2 \\ &\leq -(1 - \tilde{\beta}) \|x\|^2 + 2\tilde{\beta} \|x\| \|x^\dagger\| + \tilde{\beta} \|x^\dagger\|^2, \end{aligned}$$

because the right-hand side is bounded for all $\tilde{\beta} < 1$. So we may choose $\tilde{\beta} \in (\beta, 1)$. \square

The fact that in a linear Hilbert space setting there is always a variational source condition has been proven already in a different and more complicated way in [7, Chapter 13] based on the result that there is always a general source condition (cf. [20]). In [7, Section 13.2] it is shown that the constant β in a variational source condition does not influence the index function φ in a linear Hilbert space setting, up to some scaling.

4.2 Bregman distance in Banach spaces

A first convergence rate result for Tikhonov regularization (1.2) with convex penalty Ω can be found in [4] based on a source condition. There the error functional E^\dagger is a *Bregman distance* with respect to Ω , which is defined as follows. Let x^\dagger be an Ω -minimizing solution with nonempty subdifferential and denote by $\xi^\dagger \in X^*$ a subgradient of Ω at x^\dagger . Then

$$B_{\xi^\dagger}(x, x^\dagger) := \Omega(x) - \Omega(x^\dagger) - \langle \xi^\dagger, x - x^\dagger \rangle_{X^* \times X}, \quad x \in X,$$

is the Bregman distance between x and x^\dagger . If Ω is a Hilbert space norm, then the corresponding Bregman distance is the usual norm distance.

Bregman distances became a standard tool for convergence rate analysis in Banach spaces. The original variational source condition in [14] used them, too. With Theorem 5 we can prove the following.

Corollary 9. *Let Assumption 1 be true and assume that there is only one solution x^\dagger to (1.1) with subgradient ξ^\dagger of Ω . Let E^\dagger be the corresponding Bregman distance*

$$E^\dagger(x) = B_{\xi^\dagger}(x, x^\dagger), \quad x \in X.$$

Then there exist a constant $\beta \in (0, 1)$ and a concave index function φ such that the variational source condition (2.1) is satisfied with $M = D(F)$.

Proof. We have to show that Assumption 4 is true. Item (i) is a consequence of the weak closedness of F . Item (ii) is trivially true, because there is only one solution. Weak lower semicontinuity of

$$x \mapsto -\beta E^\dagger(x) + \Omega(x) = (1 - \beta) \Omega(x) + \beta \Omega^\dagger + \beta \langle \xi^\dagger, x - x^\dagger \rangle_{X^* \times X}$$

follows from Assumption 1 (iv) and $\beta \leq 1$.

It remains to show boundedness of $x \mapsto \tilde{\beta} E^\dagger(x) - \Omega(x)$ for some $\tilde{\beta} > \beta$. Weak sequential compactness of the sublevel sets of Ω implies boundedness of the sublevel sets. Thus, by [24, Exercise 2.41 and pages 324–326] there are a positive constant c_1 and a constant c_2 such that

$$\Omega(x) \geq c_1 \|x\| + c_2 \quad \text{for all } x \in X.$$

With this observation and with $\tilde{\beta} \in (\beta, 1]$ we obtain

$$\begin{aligned} & \tilde{\beta} E^\dagger(x) - \Omega(x) \\ &= (\tilde{\beta} - 1) \Omega(x) - \tilde{\beta} \Omega^\dagger - \tilde{\beta} \langle \xi^\dagger, x - x^\dagger \rangle_{X^* \times X} \\ &\leq (\tilde{\beta} - 1) (c_1 \|x\| + c_2) - \tilde{\beta} \Omega^\dagger + \tilde{\beta} \|\xi^\dagger\| \|x\| + \tilde{\beta} \langle \xi^\dagger, x^\dagger \rangle_{X^* \times X} \\ &= ((c_1 + \|\xi^\dagger\|) \tilde{\beta} - c_1) \|x\| + (\tilde{\beta} - 1) c_2 - \tilde{\beta} \Omega^\dagger + \tilde{\beta} \langle \xi^\dagger, x^\dagger \rangle_{X^* \times X} \end{aligned}$$

The last expression is bounded with respect to $x \in X$ if

$$\tilde{\beta} \leq \frac{c_1}{c_1 + \|\xi^\dagger\|}.$$

Thus, Assumption 4 is true for all $\beta \in (0, \frac{c_1}{c_1 + \|\xi^\dagger\|})$. □

4.3 Multiple solutions in autoconvolution

In [5] a variational source condition for complex-valued autoconvolution

$$(F(x))(s) = \int_{\max\{s-1,0\}}^{\min\{s,1\}} x(s-t)x(t) dt, \quad s \in (0,2),$$

with $X = L^2(0,1)$ and $Y = L^2(0,2)$ has been verified, if the considered exact solutions have a sparse Fourier representation.

If x^\dagger is a solution to (1.1), then $-x^\dagger$ is a second solution and there are only these two solutions. We choose Ω in (1.2) to be the L^2 -norm. There is no reason to prefer one of the two solutions and the Tikhonov minimizers will converge (subsequentially) to both solutions. Thus, for E^\dagger we choose the point-to-set distance

$$E^\dagger(x) = \text{dist}(x, \{x^\dagger, -x^\dagger\})^2 = \min\{\|x - x^\dagger\|^2, \|x + x^\dagger\|^2\}, \quad x \in X. \quad (4.1)$$

The variational source condition obtained in [5] holds only on a set M which is the union of two small balls around x^\dagger and $-x^\dagger$. The new Theorem 5 yields a variational source condition with $M = X$, but possibly with a different index function φ .

Corollary 10. *Choose X, Y and F as above. If E^\dagger is given by (4.1), then there are a constant $\beta \in (0,1)$ and a concave index function φ such that the variational source condition (2.1) is satisfied with $M = X$.*

Proof. Assumption 1 is obviously satisfied, see [1] for weak lower semicontinuity of F . Items (i) and (ii) of Assumption 4 are trivially true. For item (iii) we observe

$$E^\dagger(x) = \|x\|^2 + \|x^\dagger\|^2 - 2|\text{Re}\langle x, x^\dagger \rangle|$$

and therefore

$$-\beta E^\dagger(x) + \Omega(x) = (1 - \beta)\|x\|^2 - \beta\|x^\dagger\|^2 + 2\beta|\text{Re}\langle x, x^\dagger \rangle|,$$

which is a lower semicontinuous functional. Item (iv) follows from

$$\begin{aligned} \tilde{\beta} E^\dagger(x) - \Omega(x) &= -(1 - \tilde{\beta})\|x\|^2 - 2\tilde{\beta}|\text{Re}\langle x, x^\dagger \rangle| + \tilde{\beta}\|x^\dagger\|^2 \\ &\leq -(1 - \tilde{\beta})\|x\|^2 + 2\tilde{\beta}\|x\|\|x^\dagger\| + \tilde{\beta}\|x^\dagger\|^2 \end{aligned}$$

for $\tilde{\beta} \in (\beta, 1)$, because the last expression is bounded with respect to $x \in X$. \square

In [5] a relatively strong assumption (sparse Fourier representation) was required to obtain a variational source condition and corresponding convergence rates. With the corollary above we now have a variational source condition and rates without additional assumptions. Although we do not know φ explicitly, variational source conditions turn out to be the right tool for convergence rate analysis.

4.4 ℓ^1 -regularization

If $X = \ell^1(\mathbb{N})$ and if Ω in (1.2) is the ℓ^1 -norm, then Tikhonov regularization is also known as ℓ^1 -regularization. In [9] it was shown, that for injective bounded linear operators in (1.1) there is always a variational source condition and the index function φ can be made explicit. There, the ℓ^1 -norm distance between exact and regularized solutions was used as error functional E^\dagger , that is,

$$E^\dagger(x) = \|x - x^\dagger\|_1, \quad x \in \ell^1(\mathbb{N}).$$

This norm distance does not coincide with the Bregman distance with respect to the ℓ^1 -norm.

For noninjective bounded linear operators variational source conditions and corresponding convergence rates were derived in [8], but under additional assumptions. The proofs there are quite technical, but the interested reader easily verifies that Theorem 5 is applicable to the noninjective ℓ^1 -setting, too, if the error functional E^\dagger is the point-to-set distance

$$E^\dagger(x) = \inf_{x^\dagger \in S} \|x - x^\dagger\|_1, \quad x \in \ell^1(\mathbb{N})$$

between some $x \in \ell^1(\mathbb{N})$ and the set S of all ℓ^1 -norm minimizing solutions of (1.1). Consequently, even without injectivity we always obtain convergence rates for ℓ^1 -regularization. We do not provide a proof here, because we would have to go deep into the technicalities of [8].

4.5 Vanishing error functional

In [18] convergence rates for the Bregman distance between regularized and exact solutions are proven. But compared to our definition of the Bregman distance above the roles of regularized and exact solution are interchanged. That is, the subgradient of Ω is taken at the regularized solutions. As smoothness assumption a variational source condition (2.1) is used, but the error functional E^\dagger there is not the Bregman distance from the rate result. Instead, the choice

$$E^\dagger(x) = 0, \quad x \in X,$$

suffices to prove rates for the (twisted) Bregman distance.

With Theorem 5 we can show that this special type of variational source condition is always available.

Corollary 11. *Let Assumption 1 be true and let E^\dagger be null everywhere. Then there exists a concave index function φ such that the variational source condition (2.1) holds with $M = D(F)$.*

Proof. From Assumption 1 we easily derive Assumption 4. Thus, Theorem 5 applies. \square

Acknowledgments

The author thanks Bernd Hofmann (Chemnitz) for several valuable remarks on the first version of the manuscript.

References

- [1] S. W. Anzengruber, S. Bürger, B. Hofmann, and G. Steinmeyer. Variational regularization of complex deautoconvolution and phase retrieval in ultrashort laser pulse characterization. *Inverse Problems*, 32(3):035002 (27pp), 2016.
- [2] R. I. Boş and B. Hofmann. An extension of the variational inequality approach for nonlinear ill-posed problems. *Journal of Integral Equations and Applications*, 22(3):369–392, 2010.
- [3] M. Burger, J. Flemming, and B. Hofmann. Convergence rates in ℓ^1 -regularization if the sparsity assumption fails. *Inverse Problems*, 29:025013 (16pp), 2013.
- [4] M. Burger and S. Osher. Convergence rates of convex variational regularization. *Inverse Problems*, 20(5):1411–1421, 2004.
- [5] S. Bürger, J. Flemming, and B. Hofmann. On complex-valued deautoconvolution of compactly supported functions with sparse Fourier representation. *Inverse Problems*, 32(10):104006 (12pp), 2016.
- [6] H. W. Engl, M. Hanke, and A. Neubauer. *Regularization of Inverse Problems*. Mathematics and Its Applications. Kluwer Academic Publishers, Dordrecht, 1996.

- [7] J. Flemming. *Generalized Tikhonov regularization and modern convergence rate theory in Banach spaces*. Shaker Verlag, Aachen, 2012.
- [8] J. Flemming. Convergence rates for ℓ^1 -regularization without injectivity-type assumptions. *Inverse Problems*, 32(9):095001 (19pp), 2016.
- [9] J. Flemming and D. Gerth. Injectivity and weak*-to-weak continuity suffice for convergence rates in ℓ^1 -regularization. *Journal of Inverse and Ill-Posed Problems*, 2017. Ahead of print, available online, DOI: 10.1515/jiip-2017-0008.
- [10] J. Flemming and B. Hofmann. A new approach to source conditions in regularization with general residual term. *Numerical Functional Analysis and Optimization*, 31(3):254–284, 2010.
- [11] M. Grasmair. Generalized Bregman distances and convergence rates for non-convex regularization methods. *Inverse Problems*, 26(11):115014 (16pp), 2010.
- [12] M. Grasmair. Variational inequalities and higher order convergence rates for Tikhonov regularisation on Banach spaces. *Journal of Inverse and Ill-Posed Problems*, 21(3):379–394, 2013.
- [13] M. Grasmair, M. Haltmeier, and O. Scherzer. The residual method for regularizing ill-posed problems. *Applied Mathematics and Computation*, 218(6):2693–2710, 2011.
- [14] B. Hofmann, B. Kaltenbacher, C. Pöschl, and O. Scherzer. A convergence rates result for Tikhonov regularization in Banach spaces with non-smooth operators. *Inverse Problems*, 23(3):987–1010, 2007.
- [15] B. Hofmann and P. Mathé. Parameter choice in Banach space regularization under variational inequalities. *Inverse Problems*, 28:104006 (17pp), 2012.
- [16] T. Hohage and F. Weidling. Characterizations of variational source conditions, converse results, and maxisets of spectral regularization methods. *SIAM Journal on Numerical Analysis*, 55(2):598–620, 2017.
- [17] T. Hohage and F. Werner. Iteratively regularized Newton methods with general data misfit functionals and applications to Poisson data. *Numerische Mathematik*, 123(4):745–779, 2013.

- [18] S. Kindermann. Convex Tikhonov regularization in Banach spaces: New results on convergence rates. *Journal of Inverse and Ill-Posed Problems*, 24(3):341–350, 2016.
- [19] C. König, T. Hohage, and F. Werner. Convergence rates for exponentially ill-posed inverse problems with impulsive noise. *SIAM Journal on Numerical Analysis*, 54(1):341–360, 2016.
- [20] P. Mathé and B. Hofmann. How general are general source conditions? *Inverse Problems*, 24(1):015009 (5pp), 2008.
- [21] P. Mathé and S. V. Pereverzev. Geometry of linear ill-posed problems in variable Hilbert scales. *Inverse Problems*, 19(3):789–803, 2003.
- [22] O. Scherzer, M. Grasmair, H. Grossauer, M. Haltmeier, and F. Lenzen. *Variational Methods in Imaging*. Number 167 in Applied Mathematical Sciences. Springer, New York, 2009.
- [23] T. Schuster, B. Kaltenbacher, B. Hofmann, and K. S. Kazimierski. *Regularization Methods in Banach Spaces*, volume 10 of *Radon Series on Computational and Applied Mathematics*. De Gruyter, Berlin/Boston, 2012.
- [24] C. Zălinescu. *Convex Analysis in General Vector Spaces*. World Scientific, River Edge, 2002.