Variational smoothness assumptions in convergence rate theory—an overview

JENS FLEMMING*

October 9, 2012

Key words: nonlinear ill-posed problems, Tikhonov regularization, convergence rates, smoothness assumptions
MSC: 47J06, 47A52

Abstract

Variational smoothness assumptions are a concept for measuring abstract smoothness of solutions to operator equations. Such assumptions are useful for analyzing regularization methods, especially for proving convergence rates in Banach spaces and in more general settings. We collect results from different papers published by several authors during the last five years. The aim is to present an overview of this relatively new concept to the interested reader without going too deep into the details.

1 Ill-posed problems, regularization, convergence rates

A frequently used model for practical problems are equations

$$F(x) = y, \quad x \in \mathrm{dom}(F) \subseteq X, \quad y \in Y \tag{1.1}$$

between Banach spaces X and Y. The exact right-hand side y is typically unknown. Instead, only noisy data $y^\delta \in Y$ satisfying $\|y - y^\delta\| \leq \delta$ are available. The noise level is quantified by $\delta \geq 0$. For solving such equations numerically the mapping $F : \mathrm{dom}(F) \to Y$ has to have nice properties.

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

But often we are concerned with *ill-posed* problems causing numerical instabilities. Next to questions on existence and uniqueness of solutions the continuous dependence of solutions on the right-hand side is of great importance. In this article we assume that due to ill-posedness effects equation (1.1) cannot be solved by standard algorithms which do not take care of ill-posedness.

There are plenty of *regularization* methods yielding approximate solutions to (1.1) in a numerically stable way (see [5,23,30]). Due to its flexibility Tikhonov regularization is a popular technique. It consists in solving the minimization problem

$$T_\alpha^\delta(x) := \frac{1}{p}\|F(x) - y^\delta\|^p + \alpha\Omega(x) \to \min_{x \in \mathrm{dom}(F)}. \tag{1.2}$$

The convex and lower semi-continuous functional $\Omega : X \to (-\infty, \infty]$ stabilizes the minimization process and the regularization parameter $\alpha > 0$ controls the trade-off between data fitting and stabilization. The power $p \geq 1$ is of minor importance. It can be used to make the minimization task numerically more tractable. Although the techniques presented later in this article can be applied to other regularization methods, too, we concentrate on Tikhonov regularization.

For each parameter $\alpha > 0$ a regularization method yields an approximate solution to (1.1). Two questions arise. How to choose $\alpha = \alpha(\delta, y^\delta)$ to get a small approximation error? And, given a choice rule for α, what is the relation between the noise level and the approximation error. To the first question there exists a number of answers. In this article we touch the issue of choosing the regularization parameter only peripherally. The second question is typically answered by proving *convergence rates*. These are asymptotic upper bounds for the approximation error in terms of the noise level. Quantifying the approximation error in terms of the norm in X is straight forward. But there are only few results providing upper bounds for this case. Instead *Bregman distances* are used (see [4]). They allow to prove convergence rates and in most situations they carry enough information about the distance between regularized and exact solutions. Bregman distances are defined as follows.

Definition 1.1. Let $\bar{x} \in X$ have non-empty subdifferential $\partial\Omega(\bar{x}) := \{\bar{\xi} \in X^* : \Omega(x) \geq \Omega(\bar{x}) + \langle \bar{\xi}, x - \bar{x} \rangle\}$. The Bregman distance with respect to Ω and $\bar{\xi} \in \partial\Omega(\bar{x})$ between two elements $x \in X$ and $\bar{x} \in X$ is defined by

$$B_{\bar{\xi}}^{\Omega}(x, \bar{x}) := \Omega(x) - \Omega(\bar{x}) - \langle \bar{\xi}, x - \bar{x} \rangle.$$

If X is a Hilbert space and if $\Omega(x) = \frac{1}{2}\|x - x_0\|^2$ with fixed $x_0 \in X$ the subdifferential at $\bar{x}$ contains only the element $\bar{\xi} = \bar{x} - x_0$ and the corresponding Bregman distance is $B_{\bar{\xi}}^{\Omega}(x, \bar{x}) = \frac{1}{2}\|x - \bar{x}\|^2$. On the other hand, if $X = l^1(\mathbb{N})$ is the space of summable sequences and if $\Omega(x) := \|x\|$ the subdifferential at $\bar{x}$ may consist of many different elements. Depending on the choice of $\bar{\xi}$ the Bregman distance between two points can be zero although the two points do not coincide. In such cases Bregman distances are not suited for expressing the approximation error.

A well-known result from the theory of ill-posed problems states that without *additional assumptions* convergence rates cannot be proven (see [5, Proposition 3.11]). All ingredients of the regularization process have to play together well. The operator and the exact solutions are the main objects to be controlled. But also the stabilizing functional Ω and properties of the underlying spaces have to be incorporated into sufficient conditions for convergence rates. A widely applicable formulation of such sufficient conditions are *variational smoothness assumptions* (often also referred to as variational inequalities). This technique has been developed by different researchers during the past five years and the aim of the present article is to provide an overview of the state of the art. Up to now results are scattered over a number of papers using different notations and conventions and are therefore hardly accessible to a broader audience.

In the next section we introduce variational smoothness assumptions and derive convergence rates. Section 3 shows relations to source conditions, which are the classical concept for proving convergence rates. Approximate source conditions are an extension of source conditions which also can be related to variational smoothness assumptions, see Section 4. Some specializations in Hilbert space settings as well as some remarks on so-called converse results are collected in Section 5. Finally, in Section 6 we draw some conclusions and formulate open problems.

2 Variational smoothness assumptions (VSA)

The formal definition of a variational smoothness assumption is as follows.

Definition 2.1. Let $x^\dagger$ be a solution of (1.1) and let $\xi^\dagger \in \partial\Omega(x^\dagger)$, $M \subseteq \mathrm{dom}(F)$, and $\beta \in (0, 1]$. Further, let $\varphi : [0, \infty) \to [0, \infty)$ be a concave and non-decreasing function with $\varphi(0) = 0$ and $\lim_{t \to 0} \varphi(t) = 0$. A variational smoothness assumption is an inequality

$$\beta B_{\xi^\dagger}^{\Omega}(x, x^\dagger) \le \Omega(x) - \Omega(x^\dagger) + \varphi(\|F(x) - F(x^\dagger)\|) \tag{2.1}$$

holding for all $x \in M$.

We give some details on the components of a variational smoothness assumption.

- The constant $\beta \in (0, 1]$ has no influence on the convergence rate obtained from a variational smoothness assumption. But in some cases $\beta < 1$ allows to prove the validity of a variational smoothness assumption which would not hold for $\beta = 1$. We exclude $\beta > 1$ to guarantee convexity of $x \mapsto \beta B^{\Omega}_{\xi^\dagger}(x, x^\dagger) - \Omega(x) + \Omega(x^\dagger)$.

- The solution $x^\dagger$ is the one for which the variational smoothness assumption provides a convergence rate. Since the Bregman distance is always non-negative it is easily seen that $x^\dagger$ minimizes Ω over all solutions belonging to the set M.

- A function with the properties of φ is sometimes called concave index function. This function completely determines the convergence rate obtainable from a variational smoothness assumption. One can show that convex functions φ automatically lead to the trivial case where the exact solution $x^\dagger$ is a minimizer of Ω over the set M (see [7, Proposition 12.10] for general φ or [20, Proposition 4.3] for monomials). Thus, restriction to concave φ is reasonable.

- In the definition of variational smoothness assumptions we do not explicitly pose assumptions on the mapping F. Nonlinear as well as non-differentiable mappings can be handled.

- The set $M \subseteq \mathrm{dom}(F)$ can be regarded as the set of interesting points. To prove convergence rates M has to contain all regularized solutions x^δ_α under consideration. Given a parameter choice $(\delta, y^\delta) \mapsto \alpha(\delta, y^\delta)$ this means

$$\bigcup_{\delta \in (0,\bar{\delta}]} \bigcup_{y^\delta : \|y - y^\delta\| \leq \delta} \operatorname*{argmin}_{x \in \mathrm{dom}(F)} T^\delta_{\alpha(\delta, y^\delta)}(x) \subseteq M$$

  for some $\bar{\delta} > 0$. Typical choices are

$$M = \left\{ x \in \mathrm{dom}(F) : \tfrac{1}{p}\|F(x) - y\|^p + \bar{\alpha}\Omega(x) \leq c \right\}$$

  with fixed $\bar{\alpha} > 0$ and $c > 0$ (see [17]) or simply $M = \mathrm{dom}(F)$.

Different degrees of generality of the following convergence rate theorem can be found in [17, Theorem 4.4], [1, Theorem 4.3], [11, Corollary 3.1], [6, Theorem 2.14].

Theorem 2.2. *Let $x^\dagger$ be a solution of* (1.1) *which satisfies a variational smoothness assumption with β, $\xi^\dagger$, φ, M and denote the Tikhonov regularized solutions by x_α^δ. Then there exist a priori and a posteriori parameter choices $(\delta, y^\delta) \mapsto \alpha(\delta, y^\delta)$ such that*

$$B_{\xi^\dagger}^{\Omega}\left(x_{\alpha(\delta,y^\delta)}^\delta, x^\dagger\right) = \mathcal{O}(\varphi(\delta)) \quad \textit{if } \delta \to 0.$$

For different formulations of a priori parameter choices we refer to [11, Corollary 3.1], [6, Theorem 2.14], [19, Theorem 1]. For α choosen by the Morozow discrepancy principle the theorem is proven in [7, Theorem 4.25], [19, Theorem 2].

We close this section with remarks on the historical development of variational smoothness assumptions and on possible extensions.

In [17] the concept of variational smoothness assumptions replacing classical source conditions in Banach spaces appeared for the first time. The authors of [17] used the somewhat misleading name variational inequality in the abstract and denoted this variational formulation as source condition in a passage of the paper. But basically they did not give a name to such conditions. Later on the term variational inequalities was used by several authors for similar approaches. Only the case $\varphi(t) = ct$, $t > 0$, was considered in [17]. A step towards general φ was undertaken in [15, inequality (3.8)] where a variational smoothness assumption with monomial φ appears in a proof. Other papers dealing with monomial φ are [8, 20]. Variational smoothness assumptions with general concave φ as in the definition above appeared almost simultanously in [1, 6, 11].

Extensions of variational smoothness assumptions have been suggested in two directions. On the one hand the Bregman distance can be replaced by other functionals and on the other hand the data fitting term $x \mapsto \frac{1}{p}\|F(x) - y^\delta\|^p$ in the Tikhonov functional can be replaced by a more general fitting term. Convergence rates based on variational smoothness assumptions for error measures different from Bregman distances have been shown in [11, Corollary 3.1] and in [7, Theorem 4.11]. In both papers arbitrary error measures are applicable. Examples for non-Bregman error terms can be found for instance in [2, Lemmas 4.4 and 4.6]. Alternative data fitting functionals in combination with variational smoothness assumptions appeared for the first time in [28] (only $\varphi(t) = ct$, $c > 0$, is considered) and a deeper analysis was carried out in [6, 7]. As we will see in the succeeding sections, for standard Tikhonov regularization in Banach spaces as considered in this article variational smoothness assumptions have only a slightly broader range of applications than source conditions or approximate source conditions. But in case of non-standard data fitting terms or non-Bregman error measures convergence rates cannot be obtained via classical techniques whereas the concept of variational smoothness assumptions provides a powerful tool even in such generalized settings.

Next to Tikhonov regularization, variational smoothness assumptions yield convergence rates for iteratively regularized Newton methods (see [21, 22]) and for the residual method (see [13]).

3 VSA and source conditions

Source conditions of the form

$$\partial\Omega(x^\dagger) \cap \operatorname{ran}(F'[x^\dagger]^*) \neq \emptyset \tag{3.1}$$

are the standard tool for obtaining convergence rates in Banach spaces (see [4]). Here again $x^\dagger$ is a solution of (1.1) for which convergence rates are desired. The opertor $F'[x^\dagger]^* : Y^* \to X^*$ is the adjoint of a bounded linear operator $F'[x^\dagger] : X \to Y$ which is some kind of linearization of F in $x^\dagger$. In case of a convex or at least star-shaped (with respect to $x^\dagger$) domain $\operatorname{dom}(F)$ we can define $F'[x^\dagger]$ as a bounded linear extension of the mapping

$$x \mapsto \lim_{t\to 0} \frac{1}{t}\big(F(x^\dagger + t(x - x^\dagger)) - F(x^\dagger)\big), \quad x \in \operatorname{dom}(F).$$

If $x^\dagger$ is an interior point of $\operatorname{dom}(F)$ and F is Fréchet differentiable at $x^\dagger$ then $F'[x^\dagger]$ coincides with the Fréchet derivative. In the following we always assume that there exists such a bounded linearization $F'[x^\dagger]$.

Source conditions have the advantage that in many cases interpretations in form of differentiability of $x^\dagger$, boundary conditions, or similar properties are accessible. In contrast, interpreting variational smoothness assumptions is up to now a very difficult task. Thus, it seems to be useful to reveal the close connection between classical source conditions and certain variational smoothness assumptions, namely those with linear φ.

As a first information in this direction we cite a result from [9], which was already proven in [29, Propositions 3.35 and 3.38] under stronger assumptions.

Theorem 3.1. *Assume* $\operatorname{dom}(F) = X$ *and that there exists a linearization* $F'[x^\dagger]$ *as described above.*

(i) If $x^\dagger$ satisfies the source condition (3.1) *and the nonlinearity condition*

$$\|F'[x^\dagger](x - x^\dagger)\| \le c_{\mathrm{NL}} \|F(x) - F(x^\dagger)\| \quad \text{for all } x \in X \tag{3.2}$$

with $c_{\mathrm{NL}} \ge 0$, then there exist $\xi^\dagger \in \partial\Omega(x^\dagger)$ and $c > 0$ such that the variational smoothness assumption (2.1) *holds with $\varphi(t) = ct$, $M = X$, $\beta = 1$.*

(ii) If a variational smoothness assumption (2.1) *holds with $\varphi(t) = ct$ for some $c > 0$ and with $M = X$ and $\beta \in (0, 1]$, then $x^\dagger$ satisfies the source condition* (3.1)*.*

Note that the nonlinearity condition (3.2) can be replaced by several other assumptions on the nonlinearity structure. The theorem remains true but perhaps with a different function φ in the first part. From the theorem we deduce that in case of a bounded linear operator F we have full equivalence between source conditions and variational smoothness assumptions with linear φ and with $M = X$.

Variational smoothness assumptions with linear φ but holding only on a set M which is smaller than X can be characterized in terms of a special variant of source conditions, so-called projected source conditions.

Given a convex set $C \subseteq \mathrm{dom}(F)$ we define the normal cone $N_C(x^\dagger) := \{\xi \in X^* : \langle \xi, x - x^\dagger \rangle \le 0 \text{ for all } x \in X\}$. Assuming a projected source condition in Banach spaces means that there are $\xi^\dagger \in \partial\Omega(x^\dagger)$ and $\eta^\dagger \in Y^*$ such that

$$F'[x^\dagger]^* \eta^\dagger - \xi^\dagger \in N_C(x^\dagger). \tag{3.3}$$

The term 'projected' is used because in case of a reflexive, strictly convex, and smooth Banach space X condition (3.3) can be reformulated as

$$x^\dagger = P_C\big(x^\dagger + J_*(F'[x^\dagger]^* \eta^\dagger - \xi^\dagger)\big)$$

(see [7, page 190]). Here $P_C : X \to C$ denotes the well-defined metric projector onto the convex set C and $J_* : X^* \to X$ is the well-defined inverse of the duality mapping on X. Projected source conditions naturally appear in the context of convexly contrained Tikhonov regularization

$$\frac{1}{p}\|F(x) - y^\delta\|^p + \alpha\Omega(x) \to \min_{x \in C}.$$

First results on convexly constraint Tikhonov regularization can be found in [25]. The following theorem has been published in original form in [9] and with shortened and simplified proof in [7, Proposition 12.26 and Theorem 12.29]. It reveals a close connection between projected source conditions and variational smoothness assumptions with linear φ and $M = C$.

Theorem 3.2. *Let $C \subseteq \mathrm{dom}(F)$ be convex with $x^\dagger \in C$, assume that there exists a linearization $F'[x^\dagger]$ as described above, and define* $\mathrm{dom}(\Omega) := \{x \in X : \Omega(x) < \infty\}$.

(i) If $x^\dagger$ satisfies the projected source condition (3.3) *and the nonlinearity condition* (3.2)*, then there exist $\xi^\dagger \in \partial\Omega(x^\dagger)$ and $c > 0$ such that the variational smoothness assumption* (2.1) *holds with $\varphi(t) = ct$, $M = C$, $\beta = 1$.*

(ii) Assume $\mathrm{dom}(\Omega) \cap \mathrm{int}(C) \neq \emptyset$ *or that Ω is continuous at a point of* $\mathrm{dom}(\Omega) \cap C$*. If a variational smoothness assumption* (2.1) *holds with $\varphi(t) = ct$ for some $c > 0$ and with $M = C$ and $\beta \in (0, 1]$, then $x^\dagger$ satisfies the projected source condition* (3.1)*.*

Summarizing the results presented in this section we can say that variational smoothness assumptions with linear φ are nothing else than classical source conditions combined with a nonlinearity condition. But one should be aware of the fact, that up to now it is not clear how to express higher order source conditions in Banach spaces (see [14, 26, 27]) by variational smoothness assumptions. In case of linear operators in Hilbert spaces higher order source conditions based on the operator A^*A can be shown to be closely connected with certain variational smoothness assumptions (see [7, Chapter 13]).

4 VSA and approximate source conditions

Source conditions (3.1) in Banach spaces provide only one fixed convergence rate in the sense that either the source condition is satisfied for a solution $x^\dagger$ or not. In the first case a convergence rate can be proven and in the second no rates are obtained. To extend the range of rates the concept of approximate source conditions has been developed in [15, 16]. The idea is to measure the violation of a benchmark source condition via a distance function

$$d(r) := \inf\{\|\xi^\dagger - F'[x^\dagger]^*\eta\| : \eta \in Y^*,\ \|\eta\| \le r\}, \quad r \ge 0, \tag{4.1}$$

and to prove rates in terms of this distance function. Here again $\xi^\dagger \in \partial\Omega(x^\dagger)$ is a subgradient of Ω. One can show that the function d is non-negative, non-increasing, lower semi-continuous, and convex (see, e.g., [7, Section 12.1.3]). If the benchmark source condition (3.1) is not satisfied, then d is strictly decreasing. Convergence rates in terms of d can only be proven if $d(r) \to 0$

for $r \to \infty$, which is the case if and only if $\xi^\dagger$ belongs to the closure of $\operatorname{ran}(F'[x^\dagger]^*)$. The obtained rate is the better the faster d decays to zero at infinity.

Variational smoothness assumptions as a very flexible sufficient condition for convergence rates should cover the concept of approximate source conditions in some way. To establish a connection between these two concepts we first introduce distance functions based on variational smoothness assumptions. Then we show that there is a one-to-one correspondence between such distance functions and functions φ in variational smoothness assumptions. Finally, we present connections between both types of distance functions.

The idea to express the violation of a benchmark variational smoothness assumption in terms of a distance function has been published as approximate variational inequality or approximate variational smoothness assumption in [8]. In analogy to approximate source conditions we define the distance function

$$D_\beta(r) := \sup_{x\in M}\bigl(\beta B^\Omega_{\xi^\dagger}(x,x^\dagger)-\Omega(x)+\Omega(x^\dagger)-r\|F(x)-F(x^\dagger)\|\bigr), \quad r \geq 0, \tag{4.2}$$

which measures the violation of a variational smoothness assumption with linear φ. One can show that the function D_β is non-negative, non-increasing, lower semi-continuous, and convex (see, e.g., [7, Section 12.1.5]).

Distance functions D_β and functions φ in a variational smoothness assumption can be transformed into each other via the calculus of conjugate functions. The conjugate function $f^* : X^* \to (-\infty,\infty]$ of a lower semi-continuous and convex function $f : X^* \to (-\infty,\infty]$ is defined by

$$f^*(\xi) := \sup_{x\in X}\bigl(\langle\xi,x\rangle - f(x)\bigr), \quad \xi \in X^*,$$

and is itself lower semi-continous and convex. The following result has been published in [7, Theorem 12.32].

Theorem 4.1. *Let $x^\dagger$ be a solution of* (1.1) *and fix $\xi^\dagger \in \partial\Omega(x^\dagger)$, $\beta \in (0,1]$, and $M \subseteq \operatorname{dom}(F)$.*

(i) If the distance function D_β defined by (4.2) *decays to zero at infinity, then $x^\dagger$ satisfies a variational smoothness assumption* (2.1) *with $\varphi(t) = -D^*_\beta(-t)$.*

(ii) By Φ denote the set of all non-decreasing, concave functions φ with $\varphi(0) = 0$ and $\lim_{t\to 0}\varphi(t) = 0$ for which a variational smoothness assumption (2.1) *is satisfied and assume $\Phi \neq \emptyset$. Further assume*

$D_\beta(0) > 0$. Then the distance function D_β defined by (4.2) *is the pointwise minimum of $r \mapsto (-\varphi)^*(-r)$ over $\varphi \in \Phi$. The minimum is attained at $\varphi(t) = -D_\beta^*(-t)$.*

The theorem states that there is a one-to-one correspondence between distance functions D_β and functions φ in a variational smoothness assumption. In other words, variational smoothness assumptions and approximate variational smoothness assumptions are equivalent concepts. Consequently, for establishing a connection between variational smoothness assumptions and approximate source conditions it suffices to relate distance functions D_β to distance functions d, defined in (4.2) and (4.1), respectively. Such relations are not as strong as one might expect. The following results are collected from [1, Theorem 5.2] and [7, Proposition 12.33 and Theorem 12.35].

Theorem 4.2. *Assume that $A := F$ is bounded and linear with* $\mathrm{dom}(A) = X$ *and let the distance functions d and D_β be defined by* (4.1) *and* (4.2)*, respectively.*

(i) If M in the definition of D_β is bounded, then there is a constant $c > 0$ such that

$$D_\beta(r) \leq c d(r) \quad \textit{for all } r \geq 0.$$

(ii) If there are $q > 1$ and $c_q \geq 0$ such that

$$\frac{1}{q}\|x - x^\dagger\|^q \leq c_q B^\Omega_{\xi^\dagger}(x, x^\dagger) \quad \textit{for all } x \in M,$$

where M is from the definition of D_β, then there is a constant $c > 0$ such that

$$D_\beta(r) \leq c d(r)^{\frac{q}{q-1}} \quad \textit{for all } r \geq 0.$$

(iii) If $\beta = 1$ and if $x^\dagger$ is an interior point of M (from the definition of D_1), then there is a constant $c > 0$ such that

$$D_\beta(r) \geq c d(r) \quad \textit{for all } r \geq 0.$$

As we see from the theorem only in few cases D_β and d show the same behavior. The reason lies in the choice of the X^*-norm to measure the distance of $\xi^\dagger$ to a source condition, cf. (4.1). The norms in X and x^* have no influence on the Tikhonov functional and therefore they also do not influence the convergence rate. The next and last theorem of this section reformulates the definition of distance functions D_β in a way which shows a better measure for the distance between $\xi^\dagger$ and a source condition, namely a certain Bregman distance in X^* (compare the case $\beta = 0$ in the following theorem with (4.1)). The theorem has been published in [7, Theorem 12.37].

Theorem 4.3. *Let $A := F$ be bounded and linear with $\operatorname{dom}(A) = X$ and let $\beta \in (0,1)$, $\xi^\dagger \in \partial\Omega(x^\dagger)$, and $x^\dagger \in M \subseteq X$. Further, assume that M is convex and denote the indicator function of M by δ_M (zero on M, infinity outside M). Then $x^\dagger \in \partial(\Omega + \delta_M)(\xi^\dagger)$ and the distance function D_β defined by (4.2) is given by*

$$D_\beta(r) = (1-\beta)\inf\Big\{B_{x^\dagger}^{(\Omega+\delta_M)^*}\big(\xi^\dagger + \tfrac{1}{1-\beta}(A^*\eta - \xi^\dagger), \xi^\dagger\big) : \eta \in Y^*,\ \|\eta\| \le r\Big\}.$$

5 VSA in Hilbert spaces

In the previous two sections we have seen that in Banach spaces there is a strong connection between source conditions and variational smoothness assumptions whereas the relation between approximate source conditions and variational smoothness assumptions is comparatively weak. In the following we discuss three results in a Hilbert space setting: equivalence of approximate source conditions and variational smoothness assumptions, relations between general source conditions and variational smoothness assumptions, and a converse result.

In the present sections we assume that X and Y are Hilbert spaces and that $A : X \to Y$ is linear and bounded. By x_α^δ we denote the minimizers of the classical Tikhonov functional

$$x \mapsto \frac{1}{2}\|Ax - y^\delta\|^2 + \frac{\alpha}{2}\|x\|^2,$$

that is, $p = 2$ and $\Omega(x) = \frac{1}{2}\|x\|^2$ in (1.2). In all cases where variational smoothness assumptions or approximate variational smoothness assumptions appear we use $M := X$ for the underlying set M. Due to the linearity of A and to the structure of the Tikhonov functional this is no serious restriction.

Applying Theorem 4.3 to this Hilbert space setting we immediately obtain the following result.

Theorem 5.1. *Let D_β and d be the distance functions defined by (4.2) and (4.1), respectivley. If $\beta < 1$ then*

$$D_\beta(r) = \frac{1}{2(1-\beta)} d(r)^2 \quad \textit{for all } r \ge 0.$$

Thus, for linear operators in Hilbert spaces approximate variational smoothness assumptions (and therefore also variational smoothness assumptions) are a reformulation of approximate source conditions. Consequently, all results obtained for approximate source conditions in the literature so far also apply to (approximate) variational smoothness assumptions in Hilbert spaces.

In contrast to Banach spaces Hilbert spaces allow not only one source condition but a wide range of different *general source conditions*, which of course provide a wide range of different convergence rates. A general source condition has the form

$$x^\dagger \in \operatorname{ran}\bigl(\vartheta(A^*A)\bigr)$$

with a concave index function $\vartheta : (0, \infty) \to (0, \infty)$, that is, ϑ is concave, strictly increasing, and continuous with $\lim_{t\to 0} \vartheta(t) = 0$. The corresponding convergence rate is

$$\|x_\alpha - x^\dagger\| = \mathcal{O}(\vartheta(\alpha)) \quad \text{if } \alpha \to 0.$$

Here $x_\alpha := x_\alpha^0$ denotes the Tikhonov minimizer for exact data and $x^\dagger$ is the norm-minimizing solution to the linear equation (1.1). From the well-known estimate

$$\|x_\alpha^\delta - x^\dagger\| \leq \frac{\delta}{2\sqrt{\alpha}} + \|x_\alpha - x^\dagger\|$$

one then obtains a rate for $\|x_\alpha^\delta - x^\dagger\|$. The following theorem connects general source conditions to approximate source conditions. Proofs can be found in [18, Theorem 5.9] and [7, Theorem 13.10].

Theorem 5.2. *Let A be compact and injective and assume that $x^\dagger = \vartheta(A^*A)w$ with $\|w\| = 1$ and $x^\dagger \notin \operatorname{ran}\bigl((A^*A)^{\frac{1}{2}}\bigr)$.*

(i) If $\sigma(t) := \frac{\sqrt{t}}{\vartheta(t)}$ defines an index function, then

$$d(r) \leq r\sqrt{\sigma^{-1}\bigl(r^{-1}\bigr)}$$

for all sufficiently large r.

(ii) If ϑ^2 is concave, then

$$d(r) \leq \bigl(-\vartheta(\bullet^2)\bigr)^*(-r)$$

for all $r \geq 0$, where f^* denotes the conjugate function of f (cf. previous section).

Some calculations show that the convergence rate obtained from the estimates for the distance function d in the theorem are the same as directly obtained from a general source condition with ϑ.

Summarizing the results presented up to now we see that variational smoothness assumptions generalize classical source conditions, approximate source conditions, and also general source conditions. With respect to general source conditions it remains to answer whether there is equivalence to variational smoothness assumptions (or to approximate source conditions, which is the same in Hilbert spaces).

The answer is 'no' for the following reason. As already mentioned in [24] there is no maximal general source condition. This means that if $x^\dagger$ satisfies a general source condition with some index functions ϑ then there is always another index function which decays faster to zero if the argument goes to zero than ϑ and $x^\dagger$ satisfies a general source condition with this faster decaying function. On the other hand approximate source conditions always yield the best possible rate. This result is made precise by the following theorem which is proven in [10] in a more general version. A simplified proof of the theorem as stated here can be found in [7, Theorem 13.11].

Theorem 5.3. *Assume $d(r) > 0$ for all $r \geq 0$ and $d(r) \to 0$ if $r \to \infty$. Define $\Phi(r) := \frac{d(r)}{r}$ for $r \in (0, \infty)$. Then*

$$\tfrac{1}{2} d\big(\tfrac{3}{2}\Phi^{-1}(\sqrt{\alpha})\big) \leq \|x_\alpha - x^\dagger\| \leq 2d\big(\Phi^{-1}(\sqrt{\alpha})\big)$$

for all $\alpha > 0$.

If $d(\frac{3}{2}r) \geq cd(r)$ for some $c > 0$ the theorem states that $\|x_\alpha - x^\dagger\| \sim d\big(\Phi^{-1}(\sqrt{\alpha})\big)$. In other words, the decay of the distance function d at infinity completely determines the regularization error. The condition $d(\frac{3}{2}r) \geq cd(r)$ is satisfied if d decays not too fast, for example if $d(r) \sim r^{-a}$ for some $a > 0$.

Since from general source conditions we never obtain the best possible convergence rate, approximate source conditions (and therefore also variational smoothness assumptions) cannot be equivalent to general source conditions.

The results of this section can be extended to general benchmark source conditions, that is A^* in the definition of the distance function and A in the definition of variational smoothness assumptions are replaced by $\psi(A^*A)$ with an index function ψ. Some of the results also hold for general linear regularization schemes. The interested reader finds these things in [10] and [7, Chapter 13].

6 Conclusions and open problems

From the results summarized in this article we see that the concept of variational smoothness assumptions introduced in 2007 and extended several times during the last years generalizes many other notions of abstract smoothness, especially classical source conditions, approximate source conditions, and general source conditions. In addition variational smoothness assumptions and their extensions only mentioned but not discussed in this article allow to prove convergence rates for regularization techniques not covered by source conditions. Thus, variational smoothness assumption are a unified and powerful tool for the anaylsis of regularization methods. Their interpretation is not straight forward, but having the relations to source conditions and approximate source conditions in mind, they appear to be not as abstract as they seem to be at the first look.

Concrete interpretations without detour via other smoothness concepts are subject to present and future research. The survey of the field of applications for which source conditions do not work but variational smoothness assumptions provide convergence rates is still in progress. A first compelling example can be found in [3].

Another open problem is concerned with the fact that the convergence rate obtainable from variational smoothness assumptions in the form presented in the article is bounded by the rate $\mathcal{O}(\delta)$ for the Bregman distance. But one knows that higher rates are possible. Extending the concept of variational smoothness assumptions to cover also these higher rates is an important task for future research. A first partial result has been obtained in [12].

References

[1] R. I. Boţ 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.

[2] K. Bredies and D. A. Lorenz. Regularization with non-convex separable constraints. *Inverse Problems*, 25(8):085011 (14pp), 2009.

[3] M. Burger, J. Flemming, and B. Hofmann. Convergence rates in ℓ^1-regularization if the sparsity assumption fails. *arXiv.org*, arXiv:1209.5732v1 [math.NA], September 2012. http://arxiv.org/abs/1209.5732.

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

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

[6] J. Flemming. Theory and examples of variational regularization with non-metric fitting functionals. *Journal of Inverse and Ill-posed Problems*, 18(6):677–699, 2010.

[7] J. Flemming. *Generalized Tikhonov regularization and modern convergence rate theory in Banach spaces.* Shaker Verlag, Aachen, 2012.

[8] 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.

[9] J. Flemming and B. Hofmann. Convergence rates in constrained Tikhonov regularization: equivalence of projected source conditions and variational inequalities. *Inverse Problems*, 27(8):085001 (11pp), 2011.

[10] J. Flemming, B. Hofmann, and P. Mathé. Sharp converse results for the regularization error using distance functions. *Inverse Problems*, 27(2):025006 (18pp), 2011.

[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 improved convergence rates for Tikhonov regularization on Banach spaces. *Inverse Problems*, 2012. To appear.

[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] T. Hein. Tikhonov regularization in Banach spaces – improved convergence rates results. *Inverse Problems*, 25(3):035002 (18pp), 2009.

[15] T. Hein and B. Hofmann. Approximate source conditions for nonlinear ill-posed problems – chances and limitations. *Inverse Problems*, 25(3):035003 (16pp), 2009.

[16] B. Hofmann. Approximate source conditions in Tikhonov-Phillips regularization and consequences for inverse problems with multiplication operators. *Mathematical Methods in the Applied Sciences*, 29(3):351–371, 2006.

[17] 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.

[18] B. Hofmann and P. Mathé. Analysis of profile functions for general linear regularization methods. *SIAM Journal on Numerical Analysis*, 45(3):1122–1141, 2007.

[19] B. Hofmann and P. Mathé. Parameter choice in Banach space regularization under variational inequalities. Preprint series of the Department of Mathematics 2012-5, Chemnitz University of Technology, Chemnitz, Germany, April 2012.

[20] B. Hofmann and M. Yamamoto. On the interplay of source conditions and variational inequalities for nonlinear ill-posed problems. *Applicable Analysis*, 89(11):1705–1727, 2010.

[21] T. Hohage and F. Werner. Iteratively regularized Newton methods with general data misfit functionals and applications to Poisson data. *arXiv.org*, arXiv:1105.2690v1 [math.NA], May 2011. http://arxiv.org/abs/1105.2690.

[22] B. Kaltenbacher and B. Hofmann. Convergence rates for the iteratively regularized Gauss-Newton method in Banach spaces. *Inverse Problems*, 26(3):035007 (21pp), 2010.

[23] B. Kaltenbacher, A. Neubauer, and O. Scherzer. *Iterative Regularization Methods for Nonlinear Ill-Posed Problems*. Number 6 in Radon Series on Computational and Applied Mathematics. Walter de Gruyter, Berlin, 2008.

[24] P. Mathé and B. Hofmann. How general are general source conditions? *Inverse Problems*, 24(1):015009 (5pp), 2008.

[25] A. Neubauer. Tikhonov-regularization of ill-posed linear operator equations on closed convex sets. *Journal of Approximation Theory*, 53(3):304–320, 1988.

[26] A. Neubauer. On enhanced convergence rates for Tikhonov regularization of nonlinear ill-posed problems in Banach spaces. *Inverse Problems*, 25(6):065009 (10pp), 2009.

[27] A. Neubauer, T. Hein, B. Hofmann, S. Kindermann, and U. Tautenhahn. Improved and extended results for enhanced convergence rates of Tikhonov regularization in Banach spaces. *Applicable Analysis*, 89(11):1729–1743, 2010.

[28] C. Pöschl. *Tikhonov Regularization with General Residual Term*. PhD thesis, University of Innsbruck, Innsbruck, Austria, October 2008. Corrected version.

[29] 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.

[30] T. Schuster, B. Kaltenbacher, B. Hofmann, and K. S. Kazimierski. *Regularization Methods in Banach Spaces*. De Gruyter, Berlin, Boston, 2012.