Inverse Problems **25** (2009) 035003 (16pp) doi:10.1088/0266-5611/25/3/035003

# Approximate source conditions for nonlinear ill-posed problems—chances and limitations

**Torsten Hein and Bernd Hofmann**[1]

Department of Mathematics, Chemnitz University of Technology, D-09107 Chemnitz, Germany

E-mail: thein@mathematik.tu-chemnitz.de and hofmannb@mathematik.tu-chemnitz.de

Received 3 August 2008, in final form 5 December 2008
Published 13 January 2009
Online at stacks.iop.org/IP/25/035003

**Abstract**

In the recent past the authors, with collaborators, have published convergence rate results for regularized solutions of linear ill-posed operator equations by avoiding the usual assumption that the solutions satisfy prescribed source conditions. Instead the degree of violation of such source conditions is expressed by distance functions $d(R)$ depending on a radius $R \geqslant 0$ which is an upper bound of the norm of source elements under consideration. If $d(R)$ tends to zero as $R \to \infty$ an appropriate balancing of occurring regularization error terms yields convergence rates results. This approach was called the method of approximate source conditions, originally developed in a Hilbert space setting. The goal of this paper is to formulate chances and limitations of an application of this method to nonlinear ill-posed problems in reflexive Banach spaces and to complement the field of low order convergence rates results in nonlinear regularization theory. In particular, we are going to establish convergence rates for a variant of Tikhonov regularization. To keep structural nonlinearity conditions simple, we update the concept of degree of nonlinearity in Hilbert spaces to a Bregman distance setting in Banach spaces.

## 1. Introduction

Motivated by an idea of Baumeister's monograph (see [2, theorem 6.8]), Hofmann has developed in [12] the *method of approximate source conditions* for inverse problems with a linear forward operator mapping between infinite-dimensional Hilbert spaces. This is an approach for finding convergence rates of regularized solutions based on balancing *distance functions* that measure the degree of violation of the solution with respect to a prescribed benchmark source condition (see also [4, 6, 13, 18] for further results and [15] for an extension to general linear regularization schemes). On the other hand, Hein has emphasized in the paper [11] that for Banach spaces the missing tool of generalized source conditions exploiting index

[1] Author to whom any correspondence should be addressed.

  

functions (see [10, 24]) can also be negotiated by such distance functions. Those functions then occur in the convergence rates instead of index functions. Based on that idea we will expand the field of *low order* convergence rates results for nonlinear regularization including logarithmic rates and Hölder rates with small exponent, which seems to be rather poor up to now. On the one hand, in contrast to other authors (see, e.g., [20, 23, 29]) we abstain from using Euler–Lagrange equations of the regularization functional and hence from assuming that the solution of the nonlinear problem is an inner point of the domain of the forward operator. On the other hand, directional derivatives of the forward operator are only required in the solution point, and an interplay with the derivatives in some neighbourhood frequently exploited in the literature (see, e.g., [9, 21], and the overview of structural conditions for nonlinearities in [20]) is not required.

This paper tries to analyse the chances of an application of the *method of approximate source conditions* with distance functions as an essential tool to *nonlinear ill-posed* operator equations describing nonlinear inverse problems in *reflexive Banach spaces*. It will be shown that *structural conditions* concerning the nonlinearity of the forward operator essentially influence the success of this application (see [1, 7, 20, 21, 23] for discussions on the variety of such nonlinearity conditions in the context of convergence rates in regularization). We are going to emphasize chances of this method in the nonlinear case, but we will not conceal the limitations of such an approach that occur whenever *smallness conditions* are required. We remark that a first successful step for this method to nonlinear problems was already done in [16], however under the very specific range invariance condition.

We are going to study *convergence rates* for stable approximate solutions of *ill-posed operator equations*

$$F(u) = v \tag{1.1}$$

with an in general *nonlinear operator* $F : \mathcal{D}(F) \subseteq U \to V$ possessing the domain $\mathcal{D}(F)$ and mapping between the normed linear spaces $U$ and $V$ with norms $\|\cdot\|_U$ and $\|\cdot\|_V$, respectively. Ill-posedness denotes here the phenomenon that solutions of (1.1) need not exist for all right-hand sides $v \in V$, if they exist they need not be uniquely determined, and unfortunately small perturbations on the right-hand side may cause arbitrarily large errors in the solution. Therefore, based on *noisy data* $v^\delta \in V$ of the exact right-hand side $v$ with

$$\|v^\delta - v\|_V \leqslant \delta \tag{1.2}$$

and noise level $\delta > 0$ we consider the variant

$$T_\alpha(u) := \|F(u) - v^\delta\|_V^p + \alpha\Omega(u) \to \min \tag{1.3}$$

of *Tikhonov regularization* (see, e.g., [30] and more recently [14, 31]) using the *stabilizing functional* $\Omega : U \to [0, +\infty]$ with the domain

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

regularization parameters $\alpha > 0$ and minimizers $u_\alpha^\delta$ of (1.3). The minimization in (1.3) is subject to $u \in \mathcal{D}$ with

$$\mathcal{D} := \mathcal{D}(F) \cap \mathcal{D}(\Omega).$$

Throughout this paper we *restrict* our considerations *to the interval* $1 < p < \infty$ for the exponent in (1.3). For the case $p = 1$ we refer, for example, to [5, 14] and with respect to approximate source conditions in particular to [11].

This paper is organized as follows: in section 2 we formulate the general assumptions concerning the operators, functionals and associated spaces of the nonlinear model in this paper. In particular, for covering the structure of nonlinearity under consideration we introduce an appropriate definition of the (local) degree of nonlinearity at a solution point. Then, in section 3, for the prescribed benchmark source condition we define distance functions depending on the radius of source elements expressing the violation of that source condition. We will prove a lemma with upper bounds of a relevant dual pairing which also applies to the case that the source condition is only satisfied in an approximate manner. Based on that lemma moreover in section 3, error estimates measured by the Bregman distance and convergence rates results for Tikhonov regularized solutions are presented in a theorem for the case that the benchmark source condition is satisfied. The two main theorems of this paper can be found with proofs in section 4. They show the chances and limitations for applying the method of approximate source conditions to nonlinear operator equations by balancing the occurring distance functions. Then convergence rates up to the benchmark order can be derived when the corresponding degree of nonlinearity at the solution is a proper one. The final section 5 is devoted to some concluding remarks pointing out also two points for future work.

The authors' aim is to express the considered elements of progress in nonlinear regularization theory in a comprehensive manner and to present all the convergence rates results in a stringent form. This, however, demands the introduction of a wide field of assumptions and definitions giving this paper a rather technical outfit. In particular, we are going to extract the deficit in the occurring convergence rates when source conditions only hold in an approximate manner. The associated rate expressions seem to be bulky at first view, but we refer to remark 4.5 for interpretation and to examples 4.8 and 4.9 for illustration in cases where the explicit structure of distance functions makes the formulae more transparent.

## 2. General assumptions and the degree of nonlinearity

In order to make the results of this paper comparable to those in [14], we pose the following general assumptions, which are closely related to the assumptions in [14].

**Assumption 2.1.**

*(1) $U$ and $V$ are reflexive Banach spaces with duals $U^*$ and $V^*$, respectively. In $U$ and $V$ we consider in addition to the norm convergence the associated weak convergence. This means in $U$*

$$u_k \rightharpoonup u \quad \Longleftrightarrow \quad \langle f, u_k\rangle_{U^*,U} \to \langle f, u\rangle_{U^*,U} \qquad \forall\, f \in U^*$$

*for the dual pairing $\langle\cdot,\cdot\rangle_{U^*,U}$ with respect to $U^*$ and $U$. The weak convergence in $V$ is defined in an analogue manner.*

*(2) $F : \mathcal{D}(F) \subseteq U \to V$ is weakly continuous and $\mathcal{D}(F)$ is weakly sequentially closed, i.e.,*

$$u_k \rightharpoonup u \quad \text{in} \quad U \quad \text{with} \quad u_k \in \mathcal{D}(F) \\ \Longrightarrow \quad u \in \mathcal{D}(F) \quad \text{and} \quad F(u_k) \rightharpoonup F(u) \quad \text{in} \quad V.$$

*(3) The functional $\Omega$ is convex and weakly sequentially lower semi-continuous.*

*(4) The domain $\mathcal{D}$ is non-empty.*

*(5) For every $\alpha > 0$, $c \geqslant 0$ and $v^\delta \in V$ the sets*

$$\mathcal{M}_\alpha(c) := \{u \in \mathcal{D} : T_\alpha(u) \leqslant c\}, \tag{2.1}$$

*whenever they are non-empty, are relatively weakly sequentially compact in the following sense: every sequence $\{u_k\}$ in $\mathcal{M}_\alpha(c)$ has a subsequence, which is weakly convergent in $U$ to some element from $U$.*

As is done in numerous recent papers concerning Banach space theory of ill-posed problems, for error analysis we exploit for the functional $\Omega$ with *subdifferential* $\partial\Omega$ the *Bregman distance* $D_\xi(\cdot, u)$ of $\Omega$ at $u \in \mathcal{D}(\Omega) \subseteq U$ and at $\xi \in \partial\Omega(u) \subseteq U^*$ defined as

$$D_\xi(\tilde{u}, u) := \Omega(\tilde{u}) - \Omega(u) - \langle \xi, \tilde{u} - u \rangle_{U^*,U}, \qquad \tilde{u} \in \mathcal{D}(\Omega) \subseteq U.$$

The set

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

is called the Bregman domain.

**Example 2.2** (norm square errors)**.** Let $U$ and $V$ be Hilbert spaces. Then for the stabilizing functional

$$\Omega(u) := \|u - u^*\|_U^2 \qquad \text{with} \quad \mathcal{D}(\Omega) = U$$

we have

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

with $\mathcal{D}_B(\Omega) = U$, where the subdifferential $\partial\Omega(u)$ is a singleton everywhere characterized by the unique element $\xi = 2(u - u^*)$. For that example, the $\Omega$-minimizing solutions and the classical $u^*$-minimum norm solutions introduced in [7, 8] coincide.

**Example 2.3** ($q$-coercive Bregman distances)**.** We say, for $1 < q < \infty$, that the Bregman distance $D_\xi(\cdot, u)$ of $\Omega$ at $u \in \mathcal{D}_B(\Omega)$ and $\xi \in \partial\Omega(u)$ is *$q$-coercive* with constant $\underline{c} > 0$ if we have

$$D_\xi(\tilde{u}, u) \geqslant \underline{c}\|\tilde{u} - u\|_U^q \qquad \text{for all} \quad \tilde{u} \in \mathcal{D}(\Omega). \tag{2.2}$$

For example, according to [3, lemma 2.7] this is the case if

$$\Omega(u) := \frac{1}{q}\|u\|_U^q \tag{2.3}$$

and $U$ is a *$q$-convex* Banach space, where the geometry of reflexive Banach spaces, in general, leads to the interval $2 \leqslant q < \infty$ for the parameter $q$ (for details see, e.g., [32]).

An element $u^\dagger \in \mathcal{D}$ is called an *$\Omega$-minimizing solution* to (1.1) if

$$\Omega(u^\dagger) = \min\{\Omega(u) : F(u) = v, u \in \mathcal{D}\} < \infty.$$

Such $\Omega$-minimizing solutions exist under assumption 2.1 if (1.1) has a solution $u \in \mathcal{D}$. For the proof and further results on *existence, stability and convergence* of regularized solutions $u_\alpha^\delta$ see [14, section 3]. The requirement of [14] that $\|\cdot\|_V$ is sequentially lower semi-continuous in the weak topology is satisfied automatically here, because of our specification of the weak topology.

Now given $\delta_{\max} > 0$, we fix $\alpha_{\max} > 0$ throughout this paper and consider only *a priori* parameter choices $\alpha = \alpha(\delta)$, with $0 < \alpha(\delta) \leqslant \alpha_{\max}$, $0 < \delta \leqslant \delta_{\max}$, satisfying the sufficient convergence conditions $\alpha(\delta) \searrow 0$ and $\frac{\delta^p}{\alpha(\delta)} \to 0$ as $\delta \to 0$. Then we have by definition of $u_\alpha^\delta$,

$$\|F(u_\alpha^\delta) - v^\delta\|_V \leqslant \left[\alpha_{\max}\left(\Omega(u^\dagger) + \frac{\delta^p}{\alpha}\right)\right]^{1/p}, \qquad u_\alpha^\delta \in \mathcal{M}_{\alpha_{\max}}\left(\alpha_{\max}\left(\Omega(u^\dagger) + \frac{\delta^p}{\alpha}\right)\right).$$

Due to convergence conditions assumed above we can suppose a constant $C > 0$ such that

$$C := \sup_{0<\delta\leqslant\delta_{\max}} \frac{\delta^p}{\alpha(\delta)} < \infty \qquad \text{and we set} \qquad \rho := \alpha_{\max}(\Omega(u^\dagger) + C).$$

Thus, we have $u_\alpha^\delta, u^\dagger \in \mathcal{M}_{\alpha_{\max}}(\rho)$. Moreover, for arbitrary $u \in \mathcal{M}_{\alpha_{\max}}(\rho)$ the inequality $\|F(u) - F(u^\dagger)\|_V \leqslant \rho^{1/p} + \delta$ holds.

Now by assumption the set $\mathcal{M}_{\alpha_{\max}}(\rho)$ is *relatively weakly sequentially compact* in $U$, and every such set in a Banach space is bounded. Hence all elements of $\mathcal{M}_{\alpha_{\max}}(\rho)$ belong to a ball in $U$, and there exists a constant $0 < K_{\max} < \infty$ such that

$$\|u - u^\dagger\|_U \leqslant K_{\max} \qquad \forall u \in \mathcal{M}_{\alpha_{\max}}(\rho). \tag{2.4}$$

For our studies we need some more assumptions which are under discussion in [14] partially as special cases.

**Assumption 2.4.** *Let $F, \Omega, U, V$ and $\mathcal{D}$ satisfy assumption 2.1.*

*(1) There exists an $\Omega$-minimizing solution $u^\dagger$ which is an element of the Bregman domain $\mathcal{D}_B(\Omega)$.*

*(2) $\mathcal{D}(F)$ is starlike with respect to $u^\dagger$, that is, for every $u \in \mathcal{D}(F)$ there exists $t_0 > 0$ such that*

$$u^\dagger + t(u - u^\dagger) \in \mathcal{D}(F) \qquad \textit{for all} \quad 0 \leqslant t \leqslant t_0.$$

*(3) There is a bounded linear operator $F'(u^\dagger) : U \to V$ such that we have for the one-sided directional derivative at $u^\dagger$ and for every $u \in \mathcal{D}$ the equality*

$$\lim_{t\to 0+} \frac{1}{t}(F(u^\dagger + t(u - u^\dagger)) - F(u^\dagger)) = F'(u^\dagger)(u - u^\dagger).$$

With respect to assumption 2.4 we should note that because of the convexity of $\Omega$ the domain $\mathcal{D}$ is also starlike. The operator $F'(u^\dagger)$ has Gâteaux derivative like properties, and there is an adjoint operator $F'(u^\dagger)^* : V^* \to U^*$ defined by

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

Now it seems to be useful to update the definition of the degree of nonlinearity from [17, definition 1] to the current situation of this paper.

**Definition 2.5.** *Let $0 \leqslant c_1, c_2 \leqslant 1$ and $0 < c_1 + c_2 \leqslant 1$. We define $F$ to be nonlinear of degree $(c_1, c_2)$ for the Bregman distance $D_\xi(\cdot, u^\dagger)$ of $\Omega$ at $u^\dagger \in \mathcal{D}(F) \cap \mathcal{D}_B(\Omega) \subseteq U$ and at $\xi \in \partial\Omega(u^\dagger) \subseteq U^*$ if there is a constant $K > 0$ such that*

$$\|F(u) - F(u^\dagger) - F'(u^\dagger)(u - u^\dagger)\|_V \leqslant K\|F(u) - F(u^\dagger)\|_V^{c_1} D_\xi(u, u^\dagger)^{c_2}, \tag{2.5}$$

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

Note that the degree of nonlinearity of definition 2.5 has a local character. In short we say that $F$ is of degree $(c_1, c_2)$ at $u^\dagger$ and $\xi$ if the requirements of the above definition are satisfied.

**Remark 2.6.** Different combinations of exponents $c_1$ and $c_2$ in the (local) degree of nonlinearity characterize the variety of structural conditions imposed on the nonlinear operator $F$ in a neighbourhood of $u^\dagger$. We are going to distinguish the following cases:

(A) Case $c_1 = 1, c_2 = 0$. This first extremal case expresses a very high potential of the linear operator $F'(u^\dagger)$ to characterize the behaviour of the nonlinear operator $F$ in a neighbourhood of $u^\dagger$ (see, e.g., the discussion in [17]). Somewhat stronger than our condition of this case is the tangential cone condition (also called $\eta$-inequality, see [7, chapter 11, formula (11.6)]) being an important structural condition for the convergence rate analysis of the nonlinear Landweber method. On the other hand, by choosing $\Omega$, from example 2.2, for the Tikhonov regularization with $p = 2$ of ill-posed nonlinear equations (1.1) in Hilbert spaces $U$ and $V$ one can formulate results with *low order* Hölder rates under low order source conditions, i.e., as $\delta \to 0$ we find for all $0 < \eta \leqslant 1$ with *the a priori* parameter choice $\alpha(\delta) \sim \delta^{2/(1+\eta)}$ the rate result

$$\left\|u_\alpha^\delta - u^\dagger\right\|_U = \mathcal{O}(\delta^{\frac{\eta}{1+\eta}}), \qquad \text{when} \qquad u^\dagger - u^* = [F'(u^\dagger)^* F'(u^\dagger)]^{\eta/2}\tilde{w}, \tilde{w} \in U \quad (2.6)$$

as a consequence of theorem 1 in [17]. Namely, the sufficient condition of this theorem, which can be formulated in our notation as the inequality chain

$$c_1 \geqslant 1 + \eta(1 - c_1 - 2c_2) > 0 \tag{2.7}$$

is satisfied for all $0 < \eta \leqslant 1$ in the case (A).

With respect to the general Bregman distance setting in the Tikhonov regularization (1.3) with $p = 2$ we remark that this case (A) is sufficient for satisfying the structural assumption

$$\langle F(u) - F(u^\dagger) - F'(u^\dagger)(u - u^\dagger), w\rangle_V \leqslant \gamma \|w\|_V \|F(u) - F(u^\dagger)\|_V$$

of [5] with Banach space $U$ and Hilbert space $V$ yielding a convergence rate $D_\xi\left(u_\alpha^\delta, u^\dagger\right) = \mathcal{O}(\delta)$ under the source condition

$$\exists\, w \in V : \qquad F'(u^\dagger)^* w \in \partial\Omega(u^\dagger).$$

The case (A) was also considered in the convergence rate analysis in [11] for the exponent $p = 1$ in (1.3).

(B) Case $0 < c_1 < 1$. For this case the operator $F'(u^\dagger)$ has less than in case (A) but still enough potential to characterize $F$ in a neighbourhood of $u^\dagger$ to a certain extent. Here, for $\Omega$ from example 2.2 and for the Tikhonov regularization with $p = 2$ in Hilbert spaces $U$ and $V$, theorem 1 from [17] also applies, but (2.7) cannot hold whenever $0 < \eta < 1$. However, for $\eta = 1$ the condition (2.7) holds if and only if

$$c_1 + c_2 = 1. \tag{2.8}$$

Hence under the source conditions $u^\dagger - u^* = [F'(u^\dagger)^* F'(u^\dagger)]^{1/2}\tilde{w}$, $\tilde{w} \in U$, which are equivalent to $u^\dagger - u^* = F'(u^\dagger)^* w$, $w \in V$, in Hilbert spaces, a rate $\left\|u_\alpha^\delta - u^\dagger\right\|_U = \mathcal{O}(\delta^{1/2})$ can be found without any additional smallness condition for the combination (2.8) of exponents in the degree of nonlinearity.

(C) Case $c_1 = 0$, $c_2 = 1$. The inequality

$$\|F(u) - F(u^\dagger) - F'(u^\dagger)(u - u^\dagger)\|_V \leqslant K D_\xi(u, u^\dagger), \tag{2.9}$$

which is of interest in that second extremal case, characterizes the classical situation occurring for the Bregman distance setting in Banach spaces in [27] and for $\Omega$ from example 2.2 in Hilbert spaces occurring in [8] and [7, chapter 10]. To obtain the rate $D_\xi\left(u_\alpha^\delta, u^\dagger\right) = \mathcal{O}(\delta)$ under the source conditions $\xi = F'(u^\dagger)^* w$, $w \in V^*$, a smallness condition $K\|w\|_{V^*} < 1$ is required. The necessity of such an additional condition shows the rather loose connection between the nonlinear operator $F$ at $u^\dagger$ and its linearization $F'(u^\dagger)$.

We have excluded the situation $0 < c_2 < 1$ and $c_1 = 0$, because used techniques fail for that situation and moreover results on that case seem to be missing up to now in the literature.

## 3. Assertions for the benchmark source condition and error estimates for approximate source conditions

Only in very specific situations can it be expected that for given $\xi \in \partial\Omega(u^\dagger) \subseteq U^*$ and the $\Omega$-minimizing solution $u^\dagger \in \mathcal{D}_B(\Omega)$ a source condition

$$\xi = F'(u^\dagger)^* w, \qquad w \in V^* \tag{3.1}$$

is satisfied. In this context, (3.1) is considered in a classical way as the *benchmark source condition* of nonlinear regularization theory. However, such a source condition is always fulfilled in an approximate manner as

$$\xi = F'(u^\dagger)^* w + r, \qquad w \in V^*, \quad r \in U^*, \tag{3.2}$$

where the elements $w$ and $r$ are not determined uniquely. If we restrict the source elements to closed balls in $V^*$ with radius $R \geqslant 0$ by $\|w\|_{V^*} \leqslant R$, then we can define the *distance function* $d(R) = d_{\xi,u^\dagger}(R)$, $R \geqslant 0$ as

$$d(R) := \min_{w\in V^*:\|w\|_{V^*}\leqslant R} \|\xi - F'(u^\dagger)^* w\|_{U^*}. \tag{3.3}$$

The distance function is well defined. In particular due to the reflexivity of $V$ implying the reflexivity of $V^*$, we have for all non-negative $R$ an element $w_R \in V^*$ with $\|w_R\|_{V^*} \leqslant R$ such that, for $r_R = \xi - F'(u^\dagger)^* w_R$, the equality $\|r_R\|_{U^*} = d(R)$ gets valid ([33, section 38.3]). Obviously $d(R)$ is non-increasing. Moreover, the decay properties of the distance function measure the *degree of violation* of $\xi$ with respect to the benchmark source condition (3.1). If the source condition (3.1) is satisfied for some $w \in V^*$ with $\|w\|_{V^*} = \underline{R}$, then we have $d(R) = 0$ for $\underline{R} \leqslant R < \infty$. Otherwise $d(R)$ is strictly positive for all $0 \leqslant R < \infty$.

For an additive decomposition (3.2) of $\xi$ the following lemma can be stated:

**Lemma 3.1.** *Let $0 \leqslant c_1, c_2 \leqslant 1$ such that $0 < c_1 + c_2 \leqslant 1$ and $c_2 = 1$ if $c_1 = 0$. Moreover, let $F$ be of degree $(c_1, c_2)$ at $u^\dagger$ and $\xi$, and let the approximate source condition (3.2) hold. Then the estimate*

$$|\langle \xi, u - u^\dagger\rangle_{U^*,U}| \leqslant \beta_1 D_\xi(u, u^\dagger) + \beta_2 \|F(u) - F(u^\dagger)\|_V^\kappa + \|r\|_{U^*}\|u - u^\dagger\|_U \tag{3.4}$$

*is valid for all $u \in \mathcal{M}_{\alpha_{\max}}(\rho)$ with exponent $0 < \kappa \leqslant 1$ of the form*

$$\kappa = \begin{cases} \dfrac{c_1}{1-c_2} & \text{for} \quad 0 \leqslant c_2 < 1 \\ 1 & \text{for} \quad c_2 = 1 \end{cases}$$

*and values $\beta_1, \beta_2 \geqslant 0$ which may depend on $\|w\|_{V^*}$. In the case $c_1 > 0$ implying $0 \leqslant c_2 < 1$ the value $\beta_1$ is independent of $\|w\|_{V^*}$ and we have $\beta_1 = c_2 < 1$.*

**Proof.** We can estimate for $u \in \mathcal{M}_{\alpha_{\max}}(\rho)$

$$\begin{aligned} |\langle \xi, u - u^\dagger\rangle_{U^*,U}| &= |\langle F'(u^\dagger)^* w + r, u - u^\dagger\rangle_{U^*,U}| \\ &= |\langle w, F'(u^\dagger)(u - u^\dagger)\rangle_{V^*,V} + \langle r, u - u^\dagger\rangle_{U^*,U}| \\ &\leqslant \|w\|_{V^*}\|F'(u^\dagger)(u - u^\dagger)\|_V + \|r\|_{U^*}\|u - u^\dagger\|_U \\ &\leqslant \|w\|_{V^*}\|F(u) - F(u^\dagger) - F'(u^\dagger)(u - u^\dagger)\|_V \\ &\quad + \|w\|_{V^*}\|F(u) - F(u^\dagger)\|_V + \|r\|_{U^*}\|u - u^\dagger\|_U \\ &\leqslant K\|w\|_{V^*}\|F(u) - F(u^\dagger)\|_V^{c_1} D_\xi(u, u^\dagger)^{c_2} \\ &\quad + \|w\|_{V^*}\|F(u) - F(u^\dagger)\|_V + \|r\|_{U^*}\|u - u^\dagger\|_U. \end{aligned}$$

We recall that $\|F(u) - F(u^\dagger)\|_V \leqslant \rho^{1/p} + \delta$ for $u \in \mathcal{M}_{\alpha_{\max}}(\rho)$. Then in the cases $c_2 = 1$ and $c_1 = 0$ we have (3.4) with constants $\beta_1 = K\|w\|_{V^*}$, $\beta_2 = \|w\|_{V^*}$ and $\kappa = 1$. In the case

$c_2 = 0$ and $0 < c_1 \leqslant 1$ we have (3.4) with $\beta_1 = 0$, $\beta_2 = \|w\|_{V^*}(K + (\rho^{1/p} + \delta)^{1-c_1})$ and $\kappa = c_1$. On the other hand, for $0 < c_2 < 1$ we have $c_1 > 0$ and can exploit a variant of Young's inequality

$$ab \leqslant \varepsilon a^{p_1} + \frac{b^{p_2}}{(\varepsilon p_1)^{p_2/p_1} p_2}, \qquad a, b \geqslant 0, \quad \varepsilon > 0, \tag{3.5}$$

with conjugate exponents $p_1, p_2 > 1$ that fulfil the equality $\frac{1}{p_1} + \frac{1}{p_2} = 1$. Precisely, let $p_1 := \frac{1}{c_2}$, $p_2 := \frac{1}{1-c_2}$, $\varepsilon := \frac{1}{p_1}$, $a := D_\xi(u, u^\dagger)^{c_2}$ and $b := K\|w\|_{V^*}\|F(u) - F(u^\dagger)\|_V^{c_1}$. Then we obtain

$$K\|w\|_{V^*}\|F(u) - F(u^\dagger)\|_V^{c_1} D_\xi(u, u^\dagger)^{c_2} \leqslant c_2 D_\xi(u, u^\dagger) + (1 - c_2)(K\|w\|_{V^*})^{\frac{1}{1-c_2}} \|F(u) - F(u^\dagger)\|_V^{\frac{c_1}{1-c_2}}.$$

Thus, in general, for $c_1 > 0$ (3.4) holds with

$$0 \leqslant \beta_1 = c_2 < 1, \qquad \beta_2 = (\rho^{1/p} + \delta)^{\frac{1-c_1-c_2}{1-c_2}} \|w\|_{V^*} + (1 - c_2)K^{\frac{1}{1-c_2}} \|w\|_{V^*}^{\frac{1}{1-c_2}}, \qquad \kappa = \frac{c_1}{1 - c_2}. \tag{3.6}$$

□

**Remark 3.2.** We emphasize that the inequality (3.4) with $r = 0$, $0 \leqslant \beta_1 < 1$, $\beta_2 \geqslant 0$ and $\kappa = 1$ occurs as an assumption in [14], for which in [14, remark 4.2] sufficient conditions along the lines of assumption 2.4 were formulated.

As a first step we formulate the consequences of lemma 3.1 in the following theorem for the case $r = 0$ of fulfilled exact benchmark source condition. This theorem extends some assertions of the recent literature, in particular, from $\kappa = 1$ to the case $0 < \kappa < 1$. Its proof is formulated in analogy with the proof of theorem 4.4 in [14].

**Theorem 3.3.** *Assume that $F, \Omega, \mathcal{D}, U$ and $V$ satisfy assumptions 2.1 and 2.4. With some $0 \leqslant c_1, c_2 \leqslant 1$ such that $0 < c_1 + c_2 \leqslant 1$ and $c_2 = 1$ if $c_1 = 0$ let $F$ be of degree $(c_1, c_2)$ for the Bregman distance $D_\xi(\cdot, u^\dagger)$ of $\Omega$ at the $\Omega$-minimizing solution $u^\dagger \in \mathcal{D}_B(\Omega) \subseteq U$ of (1.1) and at $\xi \in \partial\Omega(u^\dagger) \subseteq U^*$. Furthermore, let the source condition (3.1) hold and let $\kappa$ be defined as in lemma 3.1. In the case $c_1 > 0$, we then have the convergence rate*

$$D_\xi\left(u_\alpha^\delta, u^\dagger\right) = \mathcal{O}(\delta^\kappa) \qquad \text{as} \quad \delta \to 0 \tag{3.7}$$

*for an a priori parameter choice $\alpha \asymp \delta^{p-\kappa}$. This result is also true in the alternative case $c_1 = 0$, $c_2 = 1$ when the additional smallness condition $K\|w\|_{V^*} < 1$ holds.*

**Proof.** To prove the assertion of the theorem we apply lemma 3.1 with $r = 0$ yielding for some $0 \leqslant \beta_1 < 1$ and $\beta_2 \geqslant 0$ the inequality

$$|\langle \xi, u - u^\dagger \rangle_{U^*,U}| \leqslant \beta_1 D_\xi(u, u^\dagger) + \beta_2 \|F(u) - F(u^\dagger)\|_V^\kappa, \tag{3.8}$$

for all $u \in \mathcal{M}_{\alpha_{\max}}(\rho)$. From the definition of $u_\alpha^\delta$ and (1.2) it follows that

$$\left\|F\left(u_\alpha^\delta\right) - v^\delta\right\|_V^p + \alpha D_\xi\left(u_\alpha^\delta, u^\dagger\right) \leqslant \delta^p + \alpha\left(\Omega(u^\dagger) - \Omega\left(u_\alpha^\delta\right) + D_\xi\left(u_\alpha^\delta, u^\dagger\right)\right). \tag{3.9}$$

Moreover, by the definition of the Bregman distance and by the inequality $(a + b)^\kappa \leqslant a^\kappa + b^\kappa$ for $a, b > 0$ and $0 < \kappa \leqslant 1$ we obtain that

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

Therefore, from (3.9) it follows that

$$\left\| F\left(u_\alpha^\delta\right) - v^\delta \right\|_V^p + \alpha D_\xi\left(u_\alpha^\delta, u^\dagger\right) \leqslant \delta^p + \alpha\left(\beta_1 D_\xi\left(u_\alpha^\delta, u^\dagger\right) + \beta_2\left(\left\| F\left(u_\alpha^\delta\right) - v^\delta\right\|_V^\kappa + \delta^\kappa\right)\right). \tag{3.10}$$

Using (3.5) twice with $p_1 := p/\kappa$, $p_2 := p/(p-\kappa)$, $\varepsilon = 1$, $b := \alpha\beta_2$, on the one hand with $a := \left\| F\left(u_\alpha^\delta\right) - u^\dagger\right\|_V^\kappa$ and on the other hand with $a := \delta^\kappa$, the inequalities

$$\alpha D_\xi\left(u_\alpha^\delta, u^\dagger\right) \leqslant 2\delta^p + \alpha\beta_1 D_\xi\left(u_\alpha^\delta, u^\dagger\right) + \frac{2(p-\kappa)}{(p/\kappa)^{\kappa/(p-\kappa)}p}(\alpha\beta_2)^{p/(p-\kappa)}$$

and

$$D_\xi\left(u_\alpha^\delta, u^\dagger\right) \leqslant \frac{2\delta^p + \frac{2(p-\kappa)}{(p/\kappa)^{\kappa/(p-\kappa)}p}(\alpha\beta_2)^{p/(p-\kappa)}}{\alpha\left(1-\beta_1\right)} \tag{3.11}$$

hold. This yields (3.7) for the *a priori* parameter choice $\alpha \asymp \delta^{p-\kappa}$ and proves the theorem. □

## 4. Convergence rates for approximate source conditions

In the second step we formulate as the main theorem the consequences of lemma 3.1 for the case that $\xi$ belongs to the closure of the range $\mathcal{R}(F'(u^\dagger)^*)$ of the bounded linear operator $F'(u^\dagger)^* : V^* \to U^*$ with respect to the strong norm in $U^*$ provided that $\xi$ does not fulfil the benchmark source condition (3.1) for any $w \in V^*$. This theorem complements the corresponding assertions made in [11] for $\kappa = 1$ and $p = 1$ in (1.3) to the cases $1 < p < \infty$ and $0 < \kappa < 1$.

Therefore, our focus will be now on elements $\xi \in U^*$ satisfying the condition

$$\xi \in \overline{\mathcal{R}(F'(u^\dagger)^*)}^{\|\cdot\|_{U^*}} \setminus \mathcal{R}(F'(u^\dagger)^*). \tag{4.1}$$

Then immediately from the definition of the distance function (3.3) we obtain the next lemma.

**Lemma 4.1.** *Let $\xi$ satisfy the requirement (4.1). Then the non-increasing distance function $d(R)$ is strictly positive for all $0 \leqslant R < \infty$, and it tends to zero as $R \to \infty$.*

**Remark 4.2.** We shortly discuss the strength of the assumption (4.1) in lemma 4.1.

(a) If $U$ is a Hilbert space, then with $U = \mathcal{N}(F'(u^\dagger)) \oplus \overline{\mathcal{R}(F'(u^\dagger)^*)}$ the assumption (4.1) requires that $\xi$ be orthogonal to the null-space of $F'(u^\dagger)$, which is always satisfied for an injective operator $F'(u^\dagger)$.

(b) More general an assumption $\xi \in \overline{\mathcal{R}(F'(u^\dagger)^*)}^{\|\cdot\|_{U^*}}$ is even always fulfilled for an arbitrary reflexive Banach space if $F'(u^\dagger) : U \to V$ is injective. Namely, we then have by the separation theorem $\overline{\mathcal{R}(F'(u^\dagger)^*)}^{\|\cdot\|_{U^*}} = U^*$.

(c) If $U$ is a $q$-convex Banach space with $2 \leqslant q < \infty$ and if the stabilizing functional (2.3) from example 2.3 is chosen, then we learned by [28, lemma 2.10] that at least for equations (1.1) with bounded linear operators $F : U \to V$ the subdifferential $\xi$ at any $\Omega$-minimum norm solution $u^\dagger$ satisfies the condition $\xi \in \overline{\mathcal{R}(F^*)}^{\|\cdot\|_{U^*}}$.

Now we recall the specific variant

$$\xi = F'(u^\dagger)^* w_R + r_R, \qquad w_R \in V^*, \quad \|w_R\|_{V^*} \leqslant R, \quad \|r_R\|_{U^*} = d(R), \tag{4.2}$$

of the additive decomposition (3.2) of $\xi$ for arbitrary radii $R > 0$. We assume $c_1 > 0$. This implies $0 \leqslant c_2 < 1$ and hence $\frac{1}{1-c_2} \geqslant 1$. Then taking into account (3.4) for arbitrarily fixed $R_0 > 0$ we obtain from (3.6)

$$\beta_2 \leqslant (\rho^{1/p} + \delta)^{\frac{1-c_1-c_2}{1-c_2}} R + (1-c_2) K^{\frac{1}{1-c_2}} R^{\frac{1}{1-c_2}}$$
$$\leqslant \left((\rho^{1/p} + \delta)^{\frac{1-c_1-c_2}{1-c_2}} R_0^{-\frac{c_2}{1-c_2}} + (1-c_2) K^{\frac{1}{1-c_2}}\right) R^{\frac{1}{1-c_2}},$$

for all $R \geqslant R_0 > 0$. Thus we can find a constant $0 < \tilde{K} < \infty$ independent of $R$ such that

$$|\langle \xi, u - u^\dagger \rangle_{U^*,U}| \leqslant c_2 D_\xi(u, u^\dagger) + \tilde{K} R^{\frac{1}{1-c_2}} \|F(u) - F(u^\dagger)\|_V^{\frac{c_1}{1-c_2}} + d(R)\|u - u^\dagger\|_U \tag{4.3}$$

holds for all $u \in \mathcal{M}_{\alpha_{\max}}(\rho)$ and all $R \geqslant R_0 > 0$. If $d(R) \to 0$ as $R \to \infty$ this estimate allows us to balance $R$ and $\alpha$ in an appropriate manner such that the additional term $d(R)\|u - u^\dagger\|_U$ can be handled in order to obtain error estimates of regularized solutions.

Now we are ready to formulate our first main theorem.

**Theorem 4.3.** *Assume that $F, \Omega, \mathcal{D}, U$ and $V$ satisfy assumptions 2.1 and 2.4. For some $0 < c_1 \leqslant 1, 0 \leqslant c_2 < 1$ such that $c_1 + c_2 \leqslant 1$ let $F$ be of degree $(c_1, c_2)$ for the Bregman distance $D_\xi(\cdot, u^\dagger)$ of $\Omega$ at the $\Omega$-minimizing solution $u^\dagger \in \mathcal{D}_B(\Omega) \subseteq U$ of (1.1) and at $\xi \in \partial\Omega(u^\dagger) \subseteq U^*$ satisfying the condition (4.1). Moreover, we set*

$$\kappa := \frac{c_1}{1-c_2} \tag{4.4}$$

*and introduce for $R > 0$ the functions $\Psi(R) := \frac{d(R)^{\frac{p-\kappa}{\kappa}}}{R^{\frac{p}{c_1}}}$, $\Phi(R) := \frac{d(R)^{\frac{1}{\kappa}}}{R^{\frac{1}{c_1}}}$. Then we have the convergence rate*

$$D_\xi\left(u_\alpha^\delta, u^\dagger\right) = \mathcal{O}(d(\Phi^{-1}(\delta))) \qquad \text{as} \quad \delta \to 0 \tag{4.5}$$

*when $\alpha = \alpha(\delta)$ satisfies the equation $\delta = (\alpha d(\Psi^{-1}(\alpha)))^{\frac{1}{p}}$ for sufficiently small $\delta > 0$.*

To prove this theorem we use the following lemma:

**Lemma 4.4.** *Under the assumptions of theorem 4.3 there exist constants $K_1, K_2, K_3 > 0$, for arbitrary $R > 0$, such that the estimate*

$$D_\xi\left(u_\alpha^\delta, u^\dagger\right) \leqslant K_1 \frac{\delta^p}{\alpha} + K_2 \alpha^{\frac{\kappa}{p-\kappa}} R^{\frac{p}{(p-\kappa)(1-c_2)}} + K_3 d(R) \tag{4.6}$$

*holds for all $R \geqslant R_0 > 0$ and all sufficiently small $\alpha > 0$. If, additionally, the Bregman distance is $q$-coercive in $u^\dagger$, i.e. (2.2)) holds for $u = u^\dagger$ with constant $\underline{c} > 0$ and for some $1 < q < \infty$, then we can further estimate as*

$$D_\xi\left(u_\alpha^\delta, u^\dagger\right) \leqslant \frac{q}{q-1} K_1 \frac{\delta^p}{\alpha} + \frac{q}{q-1} K_2 \alpha^{\frac{\kappa}{p-\kappa}} R^{\frac{p}{(p-\kappa)(1-c_2)}} + \left(\frac{\underline{c}^{-\frac{1}{q}}}{1-c_2}\right)^{\frac{q}{q-1}} d(R)^{\frac{q}{q-1}}. \tag{4.7}$$

**Proof.** By definition we have $0 \leqslant \beta_1 = c_2 < 1$. Following the lines of the proof of theorem 3.3 by using inequality (4.3) instead of (3.8) we arrive at

$$D_\xi\left(u_\alpha^\delta, u^\dagger\right) \leqslant \frac{2}{1-c_2} \frac{\delta^p}{\alpha} + \frac{2(p-\kappa)}{(p/\kappa)^{\kappa/(p-\kappa)} p(1-c_2)} \alpha^{\frac{\kappa}{p-\kappa}} \beta_2^{\frac{p}{p-\kappa}} + \frac{K_{\max}}{1-c_2} d(R).$$

Here, we additionally used $\left\|u_\alpha^\delta - u^\dagger\right\|_U \leqslant K_{\max}$. Moreover, from the considerations above we have $\beta_2^{\frac{p}{p-\kappa}} \leqslant \tilde{K}^{\frac{p}{p-\kappa}} R^{\frac{p}{p-\kappa}}$. Thus, estimate (4.6) holds with $K_1 := \frac{2}{1-c_2}$, $K_2 := \frac{2(p-\kappa)}{(p/\kappa)^{\kappa/(p-\kappa)} p(1-c_2)} \tilde{K}^{\frac{p}{p-\kappa}}$ and $K_3 := \frac{K_{\max}}{1-c_2}$.

Under the additional $q$-coercivity condition (2.2) we can further conclude

$$D_\xi\left(u_\alpha^\delta, u^\dagger\right) \leqslant K_1\frac{\delta^p}{\alpha} + K_2\alpha^{\frac{\kappa}{p-\kappa}}R^{\frac{p}{(p-\kappa)(1-c_2)}} + \frac{\left\|u_\alpha^\delta - u^\dagger\right\|_U}{1-c_2}d(R)$$
$$\leqslant K_1\frac{\delta^p}{\alpha} + K_2\alpha^{\frac{\kappa}{p-\kappa}}R^{\frac{p}{(p-\kappa)(1-c_2)}} + \frac{\underline{c}^{-\frac{1}{q}}}{1-c_2}D_\xi\left(u_\alpha^\delta, u^\dagger\right)^{\frac{1}{q}}d(R).$$

Now we apply once more the inequality (3.5) with $p_1 := q$, $p_2 := q/(q-1)$, $a := D_\xi\left(u_\alpha^\delta, u^\dagger\right)^{\frac{1}{q}}$ and $b := \frac{\underline{c}^{-\frac{1}{q}}}{1-c_2}d(R)$. This yields

$$\frac{\underline{c}^{-\frac{1}{q}}}{1-c_2}D_\xi\left(u_\alpha^\delta, u^\dagger\right)^{\frac{1}{q}}d(R) \leqslant \frac{1}{q}D_\xi\left(u_\alpha^\delta, u^\dagger\right) + \left(\frac{q-1}{q}\right)\left(\frac{\underline{c}^{-\frac{1}{q}}}{1-c_2}\right)^{\frac{q}{q-1}} d(R)^{\frac{q}{q-1}},$$

and hence

$$\left(1-\frac{1}{q}\right)D_\xi\left(u_\alpha^\delta, u^\dagger\right) \leqslant K_1\frac{\delta^p}{\alpha} + K_2\alpha^{\frac{\kappa}{(p-\kappa)(1-c_2)}}R^{\frac{p}{p-\kappa}} + \left(\frac{q-1}{q}\right)\left(\frac{\underline{c}^{-\frac{1}{q}}}{1-c_2}\right)^{\frac{q}{q-1}} d(R)^{\frac{q}{q-1}}$$

completing the proof of the lemma. □

**Proof of theorem 4.3.** Now we complete the proof of the theorem based on the result of lemma 4.4. First we note that all exponents occurring in the functions $\Phi$ and $\Psi$ are strictly positive. Since $d(R)$ is non-increasing and tends to zero as $R \to \infty$ because of lemma 4.1, it is an immediate consequence that both functions $\Phi(R)$ and $\Psi(R)$ are strictly decreasing for all $R > 0$ and tend to zero as $R \to \infty$. Furthermore, the inverse functions $\Phi^{-1}$ and $\Psi^{-1}$ are well defined and strictly decreasing for sufficiently small positive arguments. We can balance now in (4.6) the last two terms as $d(R) = \alpha^{\frac{\kappa}{p-\kappa}}R^{\frac{p}{(p-\kappa)(1-c_2)}}$ or equivalently $\Psi(R) = \frac{d(R)^{\frac{p-\kappa}{\kappa}}}{R^{\frac{p}{c_1}}} = \alpha$. Evidently, we find for sufficiently small $\alpha > 0$ a uniquely determined $R = R(\alpha) > 0$ satisfying the equation $\Psi(R) = \alpha$, where $R(\alpha)$ tends to infinity as $\alpha \to 0$. Thus we can estimate further with some more constant $K_0 > 0$ and for sufficiently small $\alpha > 0$ as

$$D_\xi\left(u_\alpha^\delta, u^\dagger\right) \leqslant K_1\frac{\delta^p}{\alpha} + K_0 d(\Psi^{-1}(\alpha)).$$

The function $d(\Psi^{-1}(\alpha))$ in the last term of that estimate defined for sufficiently small $\alpha > 0$ is strictly increasing and tends to zero as $\alpha \to 0$.

In the last step we have to balance $\alpha$ and $\delta$ in the sense of an *a priori* parameter choice $\alpha = \alpha(\delta)$. For $\delta > 0$ sufficiently small, this can be done by choosing $\alpha > 0$ such that

$$\delta = (\alpha d(\Psi^{-1}(\alpha)))^{\frac{1}{p}} = \left(\frac{d(R)^{\frac{p}{\kappa}}}{R^{\frac{p}{c_1}}}\right)^{\frac{1}{p}} = \Phi(R).$$

Hence we obtain the rate (4.5). □

**Remark 4.5.** In order to interpret the convergence rate (4.5) and to compare it with the rate (3.7) occurring in the case that the benchmark source condition (3.1) is satisfied, we introduce the quotient function

$$\zeta(\delta) := \frac{\delta^\kappa}{d(\Phi^{-1}(\delta))} \tag{4.8}$$

defined for sufficiently small $\delta > 0$. Following the steps of the proof of theorem 4.3 presented above we find with $\delta = \Phi(R)$ the equations

$$\zeta(\delta) = \frac{\Phi(R)^\kappa}{d(R)} = \left(\frac{d(R)^{\frac{1}{\kappa}}}{R^{\frac{1}{c_1}}}\right)^\kappa \frac{1}{d(R)} = R^{-\frac{\kappa}{c_1}} = \left(\Phi^{-1}(\delta)\right)^{-\frac{\kappa}{c_1}} \to 0 \qquad \text{as} \quad \delta \to 0.$$

Hence there is a deficit in the convergence rate expressed by the function $\zeta$ coming from the violation of the benchmark source condition. The slower the distance function $d(R)$ declines to zero as $R \to \infty$ the greater is the deficit. For illustration we refer to example 4.9.

If the Bregman distance is $q$-coercive we are able to present another result with improved convergence rates. This is done in the second main theorem.

**Theorem 4.6.** *Let the assumptions of theorem 4.3 hold including the setting (4.4) of $\kappa$ and let the Bregman distance be q-coercive in $u^\dagger$ with constant $\underline{c} > 0$ and for some $1 < q < \infty$. Moreover, we introduce for $R > 0$ the functions $\Psi_q(R) := \frac{d(R)^{\frac{q(p-\kappa)}{(q-1)\kappa}}}{R^{\frac{p}{c_1}}}$ and $\Phi_q(R) := \frac{d(R)^{\frac{q}{(q-1)\kappa}}}{R^{\frac{1}{c_1}}}$. Then we have the convergence rate*

$$D_\xi\left(u_\alpha^\delta, u^\dagger\right) = \mathcal{O}\left(d\left(\Phi_q^{-1}(\delta)\right)^{\frac{q}{q-1}}\right) \qquad \text{as} \quad \delta \to 0 \tag{4.9}$$

*when $\alpha = \alpha(\delta)$ satisfies the equation $\delta = \alpha^{\frac{1}{p}} d\left(\Psi_q^{-1}(\alpha)\right)^{\frac{q}{p(q-1)}}$ for sufficiently small $\delta > 0$.*

**Proof.** We consider the estimate (4.7) and balance the last two terms on the right-hand side as an appropriate one-to-one correspondence between sufficiently large $R$ and sufficiently small $\alpha > 0$,

$$\alpha^{\frac{\kappa}{p-\kappa}} R^{\frac{p}{(p-\kappa)(1-c_2)}} = d(R)^{\frac{q}{q-1}}, \qquad \text{or equivalently} \qquad \Psi_q(R) = \frac{d(R)^{\frac{q(p-\kappa)}{(q-1)\kappa}}}{R^{\frac{p}{c_1}}} = \alpha.$$

In the second step we equilibrate the remaining terms in $\delta$ and $\alpha$ as

$$d(R)^{\frac{q}{q-1}} = \frac{\delta^p}{\alpha} = \frac{\delta^p}{d(R)^{\frac{q(p-\kappa)}{(q-1)\kappa}}} R^{\frac{p}{c_1}},$$

which gives

$$\delta^p R^{\frac{p}{c_1}} = d(R)^{\frac{q}{q-1}(1+\frac{p-\kappa}{\kappa})} = d(R)^{\frac{qp}{(q-1)\kappa}}$$

and yields with $\Phi_q(R) = d(R)^{q/((q-1)\kappa)} R^{(-1/c_1)} = \delta$ and

$$\delta = \alpha^{\frac{1}{p}} d\left(\Psi_q^{-1}(\alpha)\right)^{\frac{q}{(q-1)p}} = \alpha^{\frac{1}{p}} d(R)^{\frac{q}{(q-1)p}}$$

the estimate (4.9) when the corresponding *a priori* parameter choice is taken. This proves the theorem. □

The $q$-coercivity of the Bregman distance also allows us to derive convergence rates with respect to the norm $\|\cdot\|_U$ instead of $D_\xi\left(u_\alpha^\delta, u^\dagger\right)$. This is an immediate consequence of the formulae (4.9) and (2.2).

**Corollary 4.7.** *Under the conditions and notations of theorem 4.6 we have the convergence rate*

$$\left\|u_\alpha^\delta - u^\dagger\right\|_U = \mathcal{O}\left(d\left(\Phi_q^{-1}(\delta)\right)^{\frac{1}{q-1}}\right) \qquad \text{as} \quad \delta \to 0. \tag{4.10}$$

There are different ways to obtain distance functions $d(R)$ or at least appropriate majorants based on *link conditions*, for example *range inclusions*, between the linearization operator

$F'(u^\dagger)$ and the self-adjoint positive operators $G : U \to U$ that express well known or assumed smoothness properties of the solution in the form $\xi \in \mathcal{R}(G)$ (see [4, 15, 18]).

In [12, p 358–9] both situations of slow logarithmic decay rates and of faster power decay rates of $d(R)$ were discussed in a more simpler setting. Below we outline the consequences of these situations in two examples.

**Example 4.8** (logarithmic convergence rates)**.** First we consider a slow decay rate

$$d(R) \leqslant \frac{C}{(\log R)^\mu}$$

of logarithmic type for the distance function considered for sufficiently large $R > 0$ and some exponent $\mu > 0$. This expresses the fact that $\xi$ violates the benchmark source condition (3.1) in a strong manner. If $U$ was a Hilbert space, then one should expect that $\xi$ would satisfy only a logarithmic source condition (cf [19])

$$\xi = \varphi(F'(u^\dagger)^* F'(u^\dagger))w, \quad w \in U, \quad \text{with} \quad \varphi(t) = 1/(\log(1/t))^\mu.$$

Now in our Banach spaceworld we apply the estimate (4.6) from lemma 4.4. We set $R := \alpha^{-\nu}$ with $0 < \nu < \frac{\kappa(1-c_2)}{p}$ to obtain an estimate

$$D_\xi\left(u_\alpha^\delta, u^\dagger\right) \leqslant K_1 \frac{\delta^p}{\alpha} + \frac{K_4}{(\log(1/\alpha))^\mu}$$

for sufficiently small $\alpha > 0$. Here we used that for small $\alpha$ the logarithmic rate is slower than any power rate with positive exponent. Now for any *a priori* parameter choice $\alpha \asymp \delta^\gamma$ with exponent $0 < \gamma < p$ in the sense of theorem 4.3 we arrive at the convergence rate

$$D_\xi\left(u_\alpha^\delta, u^\dagger\right) = \mathcal{O}\left(\frac{1}{(\log(1/\delta))^\mu}\right) \qquad \text{as} \quad \delta \to 0. \tag{4.11}$$

If the additional condition (2.2) is valid and we estimate in the sense of theorem 4.6 starting with the estimate (4.7), then we arrive at the logarithmic rate

$$D_\xi\left(u_\alpha^\delta, u^\dagger\right) = \mathcal{O}\left(\frac{1}{(\log(1/\delta))^{\frac{\mu q}{q-1}}}\right) \qquad \text{as} \quad \delta \to 0. \tag{4.12}$$

For all $2 \leqslant q < \infty$, the convergence rate (4.12) is better than the rate (4.11).

**Example 4.9** (Hölder convergence rates)**.** As a second situation we consider a power-type decay rate

$$d(R) \leqslant \frac{C}{R^{\frac{\mu}{1-\mu}}}, \qquad 0 < \mu < 1, \tag{4.13}$$

for the distance function considered for sufficiently large $R > 0$. This expresses the fact that $\xi$ violates the benchmark source condition (3.1) in a medium manner. If the parameter $\mu$ varies through the range $0 < \mu < 1$, then all possible powers $R^\theta$, $0 < \theta < \infty$, occur in the denominator of the right-hand side of (4.13).

If $U$, $V$ were Hilbert spaces, then under the situation of example 2.2 by the converse result of corollary 3.3 in [6] the decay rate (4.13) of the distance function would imply the range condition $u^\dagger - u^* \in \mathcal{R}((F'(u^\dagger)^* F'(u^\dagger))^{\nu/2})$ for all $0 < \nu < \mu$. Then with $c_1 = 1, c_2 = 0$ we would also find low order convergence rates $\left\|u_\alpha^\delta - u^\dagger\right\|_U = \mathcal{O}(\delta^{\frac{\nu}{1+\nu}})$, for all $0 < \nu < \mu$. However, for $0 < c_1 < 1$, no such results would come from [17, theorem 1]. Noting that by [6, theorem 3.1] on the other hand a low order source condition $u^\dagger - u^* \in \mathcal{R}((F'(u^\dagger)^* F'(u^\dagger))^{\mu/2})$ would imply a decay rate of type (4.13) for the distance function, we can compare now the Hilbert space results with the convergence rates obtained by theorems 4.3 and 4.6.

For the Bregman distance and Banach space setting we find with (4.13) from theorem 4.3 a function $\Phi(R) = CR^{\frac{1-\mu c_2}{(\mu-1)c_1}}$ in the case $0 < c_1 \leqslant 1$ with $0 < c_1 + c_2 \leqslant 1$, which yields with $\kappa = \frac{c_1}{1-c_2}$ the convergence rate

$$D_\xi\left(u_\alpha^\delta, u^\dagger\right) = \mathcal{O}\left(\delta^{\frac{\mu c_1}{1-\mu c_2}}\right) = \mathcal{O}\left(\delta^{\kappa\left(\frac{\mu-\mu c_2}{1-\mu c_2}\right)}\right) \qquad \text{as} \quad \delta \to 0. \tag{4.14}$$

Even if $\kappa = \frac{c_1}{1-c_2} = 1$, i.e. $c_1 + c_2 = 1$, the corresponding Hölder rate exponent $0 < \frac{\mu c_1}{\mu c_1 + (1-\mu)} < 1$ tends to zero as $c_1 \to 0$, whereas the rate $\delta^\kappa$ of theorem 3.3 remains valid for arbitrarily small $c_1 > 0$.

The behaviour of the function (4.8) introduced in remark 4.5 can be illustrated in this example by the explicit order expression

$$\zeta(\delta) = (\Phi^{-1}(\delta))^{\frac{\kappa}{c_1}} \sim \delta^{\kappa\left(\frac{1-\mu}{1-\mu c_2}\right)}.$$

This function characterizes the rate deficit caused by violating the benchmark source condition, where the deficit grows with growing exponents of $\delta$ and $\zeta(\delta) \sim \delta^\kappa$ would express the limiting worst case. Now the exponent grows with the amplification factor $\frac{1-\mu}{1-\mu c_2}$ which increases for fixed $c_2$ when $\mu$ decreases. The slower the distance function $d(R)$ declines to zero as $R \to \infty$ the greater is the deficit.

On the other hand, under the additional $q$-coercivity condition (2.2) for the Bregman distance theorem 4.6 yields the rate

$$D_\xi\left(u_\alpha^\delta, u^\dagger\right) = \mathcal{O}\left(\delta^{\frac{\mu c_1 q}{(1-\mu c_2)q+\mu-1}}\right) \qquad \text{as} \quad \delta \to 0, \tag{4.15}$$

which is better than (4.14). Provided that the situation of example 2.2 arises we have $q = 2$, and (4.15) gives here the rate

$$\left\|u_\alpha^\delta - u^\dagger\right\|_U = \mathcal{O}\left(\delta^{\frac{\mu c_1}{1+\mu(1-2c_2)}}\right) = \mathcal{O}\left(\delta^{\frac{\mu c_1}{(1-\mu c_2)+\mu(1-c_2)}}\right) \qquad \text{as} \quad \delta \to 0.$$

Evidently, the low order rate results of theorems 4.3 and 4.6 are more general than the older ones in Hilbert space, because they include the variation of all three parameters $c_1, c_2$ and $\mu$. In particular, theorem 4.6 even leads to optimal convergence rate $\left\|u_\alpha^\delta - u^\dagger\right\|_U = \mathcal{O}\left(\delta^{\frac{\mu}{1+\mu}}\right)$ for $c_1 = 1, c_2 = 0$ in the case of example 2.2.

**Remark 4.10.** In the process of balancing the distance functions $d(R)$ the proofs of theorems 4.3 and 4.6 both exploited the estimate (4.3) with the last term $d(R)\|u - u^\dagger\|_U$ on the right-hand side. The proof of the latter theorem used the fact that $\|u - u^\dagger\|_U$ tends to zero for $u := u_\alpha^\delta$ and $\alpha(\delta) \to 0$, whereas the proof of the former only used the boundedness (2.4). So it is not amazing that the rate (4.9) tends to be better than the rate (4.5) as the examples 4.8 and 4.9 show. However, one should note that the $q$-coercivity (2.2) in theorem 4.6 may be a strong additional requirement.

## 5. Conclusions

As the proofs of lemma 4.4 and of theorem 4.3 show, for the applicability of the method of approximate source conditions based on balancing large $R$ and small $\alpha$ as developed for linear ill-posed problems in [12, 13, 15] to nonlinear ill-posed problems the exponent $c_1$ in definition 2.5 has to be strictly positive. Only for such cases we have automatically $\beta_1 < 1$ in the estimate (3.4) which allows us to use arbitrarily large values $R$. In the alternative case $c_1 = 0$ and $c_2 = 1$ as in the classical theory of [8] an additional smallness condition

$$K\|w\|_{V^*} = KR < 1$$

is required that restricts the radii $R$ by $R \leqslant R_{\max} < \infty$ for ensuring $\beta_1 < 1$. This restriction destroys the success of the balancing approach, and hence convergence rates cannot be derived in such a way. It is a simple consequence of our ideas using (3.1) as the benchmark source condition that we can present here only results with *low order convergence rates*, i.e., the rate is not better than $D_\xi\left(u_\alpha^\delta, u^\dagger\right) = \mathcal{O}(\delta)$. However, we cannot answer the question of whether for nonlinear ill-posed equations the method of approximate source conditions may yield higher convergence rates when benchmark source conditions with more smoothness are exploited. For linear ill-posed equations an extension of the method to general index functions as a benchmark was successful (see [6, 15]). On the other hand, as a rule faster convergence rates for nonlinear ill-posed equations require additional conditions associated with smallness (see for details [25] and [22, 26, 29]).

Finally let us mention two points for future work. First, in any case the abstract theory presented here has to be complemented by illustrative examples with concrete nonlinear forward operators and concrete Banach spaces. In particular, examples with fractional exponents $0 < c_1 < 1$ would be of interest. Second, we assumed $c_1 + c_2 \leqslant 1$ throughout this paper. It is open whether the case $c_1 = 1, c_2 > 0$ can lead to further convergence rates results. At least, for example 2.2, in Hilbert spaces $U$ and $V$ inequalities of the form

$$\|F(u) - F(u^\dagger) - F'(u^\dagger)(u - u^\dagger)\|_V \leqslant K\|F(u) - F(u^\dagger)\|_V\|u - u^\dagger\|_U$$

have numerous applications which correspond to $c_1 = 1, c_2 = 1/2$ in our notation.

## Acknowledgments

Both authors express their thanks to two anonymous referees for kind suggestions that helped us substantially to improve the text, and to Dr Radu Ioan Boţ (TU Chemnitz) for fruitful discussions and hints concerning the functional analytic background of this paper. Moreover, the second author acknowledges support by Deutsche Forschungsgemeinschaft (DFG) under Grant HO 1454/7-2.

## References

[1] Bakushinsky A B and Kokurin M Yu 2004 *Iterative Methods for Approximate Solutions of Inverse Problems* (Dordrecht: Springer)

[2] Baumeister J 1987 *Stable Solution of Inverse Problems* (Braunschweig: Vieweg)

[3] Bonesky T, Kazimierski K S, Maass P, Schöpfer F and Schuster T 2008 Minimization of Tikhonov functionals in Banach spaces *Abstr. Appl. Anal.* (DOI:10.1155/2008/192679) 19

[4] Böttcher A, Hofmann B, Tautenhahn U and Yamamoto M 2006 Convergence rates for Tikhonov regularization from different kinds of smoothness conditions *Appl. Anal.* **85** 555–78

[5] Burger M and Osher S 2004 Convergence rates of convex variational regularization *Inverse Problems* **20** 1411–21

[6] Düvelmeyer D, Hofmann B and Yamamoto M 2007 Range inclusions and approximate source conditions with general benchmark functions *Numer. Funct. Anal. Optim.* **28** 1245–61

[7] Engl H W, Hanke M and Neubauer A 1996 *Regularization of Inverse Problems* (Dordrecht: Kluwer)

[8] Engl H W, Kunisch K and Neubauer A 1989 Convergence rates for Tikhonov regularization of nonlinear ill-posed problems *Inverse Problems* **5** 523–40

[9] Hanke M, Neubauer A and Scherzer O 1995 A convergence analysis of the Landweber iteration for nonlinear ill-posed problems *Numer. Math.* **72** 21–37

[10] Hegland M 1995 Variable Hilbert scales and their interpolation inequalities with applications to Tikhonov regularization *Appl. Anal.* **59** 207–23

[11] Hein T 2008 Convergence rates for regularization of ill-posed problems in Banach spaces by approximate source conditions *Inverse Problems* **24** 045007

[12] Hofmann B 2006 Approximate source conditions in Tikhonov–Phillips regularization and consequences for inverse problems with multiplication operators *Math. Methods Appl. Sci.* **29** 351–71

[13] Hofmann B, Düvelmeyer D and Krumbiegel K 2006 Approximate source conditions in regularization—new analytical results and numerical studies *J. Math. Mod. Anal.* **11** 41–56

[14] Hofmann B, Kaltenbacher B, Pöschl C and Scherzer O 2007 A convergence rates result for Tikhonov regularization in Banach spaces with non-smooth operators *Inverse Problems* **23** 987–1010

[15] Hofmann B and Mathé P 2007 Analysis of profile functions for general linear regularization methods *SIAM J. Numer. Anal.* **45** 1122–41

[16] Hofmann B, Mathé P and Pereverzev S V 2007 Regularization by projection: approximation theoretic aspects and distance functions *J. Inv. Ill-Posed Problems* **15** 527–45

[17] Hofmann B and Scherzer O 1994 Factors influencing the ill-posedness of nonlinear problems *Inverse Problems* **10** 1277–97

[18] Hofmann B and Yamamoto M 2005 Convergence rates for Tikhonov regularization based on range inclusions *Inverse Problems* **21** 805–20

[19] Hohage T 1999 Iterative methods in inverse obstacle scattering: regularization theory of linear and nonlinear exponentially ill-posed problems *PhD Thesis University of Linz, Austria*

[20] Kaltenbacher B 2008 A note on logarithmic convergence rates for nonlinear Tikhonov regularization *J. Inv. Ill-Posed Problems* **16** 79–88

[21] Kaltenbacher B, Neubauer A and Scherzer O 2008 *Iterative Regularization Methods for Nonlinear Ill-Posed Problems* (Berlin: de Gruyter & Co)

[22] Lu S, Pereverzev S V and Ramlau R 2007 An analysis of Tikhonov regularization for nonlinear ill-posed problems under a general smoothness assumption *Inverse Problems* **23** 217–30

[23] Mahale P and Nair M T 2007 Tikhonov regularization of nonlinear ill-posed equations under general source conditions *J. Inv. Ill-Posed Problems* **15** 813–29

[24] Mathé P and Pereverzev S V 2003 Geometry of linear ill-posed problems in variable Hilbert scales *Inverse Problems* **19** 789–803

[25] Neubauer A 1989 Tikhonov regularisation for non-linear ill-posed problems: optimal convergence rates and finite-dimensional approximation *Inverse Problems* **5** 541–57

[26] Resmerita E 2005 Regularization of ill-posed problems in Banach spaces: convergence rates *Inverse Problems* **21** 1303–14

[27] Resmerita E and Scherzer O 2006 Error estimates for non-quadratic regularization and the relation to enhancement *Inverse Problems* **22** 801–14

[28] Schöpfer F, Louis A K and Schuster T 2006 Nonlinear iterative methods for linear ill-posed problems in Banach spaces *Inverse Problems* **22** 311–29

[29] Tautenhahn U and Jin Q 2003 Tikhonov regularization and a posteriori rules for solving nonlinear ill-posed problems *Inverse problems* **19** 1–21

[30] Tikhonov A N and Arsenin V Y 1977 *Solutions of Ill-Posed Problems* (New York: Wiley)

[31] Vasin V V and Korotkii M A 2007 Tikhonov regularization with nondifferentiable stabilizing functionals *J. Inv. Ill-Posed Problems* **15** 853–65

[32] Xu Z-B and Roach G 1991 Characteristic inequalities of uniformly convex and uniformly smooth Banach spaces *J. Math. Anal. Appl.* **157** 189–210

[33] Zeidler E 1985 *Nonlinear Functional Analysis and its Applications: III. Variational Methods and Optimization* (Berlin : Springer)