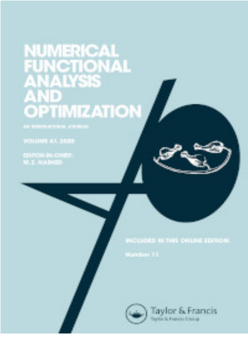

Numerical Functional Analysis and Optimization

ISSN: 0163-0563 (Print) 1532-2467 (Online) Journal homepage: https://www.tandfonline.com/loi/lnfa20

Source Conditions for Non-Quadratic Tikhonov Regularization

Markus Grasmair

To cite this article: Markus Grasmair (2020) Source Conditions for Non-Quadratic Tikhonov Regularization, Numerical Functional Analysis and Optimization, 41:11, 1352-1372, DOI: 10.1080/01630563.2020.1772289

To link to this article: https://doi.org/10.1080/01630563.2020.1772289

Published online: 12 Jun 2020.

NUMERICAL FUNCTIONAL ANALYSIS AND OPTIMIZATION
2020, VOL. 41, NO. 11, 1352–1372
https://doi.org/10.1080/01630563.2020.1772289

Source Conditions for Non-Quadratic Tikhonov Regularization

Markus Grasmair

Department of Mathematical Sciences, Norwegian University of Science and Technology, Trondheim, Norway

ABSTRACT
In this paper, we consider convex Tikhonov regularization for the solution of linear operator equations on Hilbert spaces. We show that standard fractional source conditions can be employed in order to derive convergence rates in terms of the Bregman distance, assuming some stronger convexity properties of either the regularization term or its convex conjugate. In the special case of quadratic regularization, we are able to reproduce the whole range of Hölder type convergence rates known from classical theory.

ARTICLE HISTORY
Received 24 February 2020
Revised 9 May 2020
Accepted 18 May 2020

KEYWORDS
convergence rates; Linear inverse problems; source conditions; Tikhonov regularization

2010 MATHEMATICS SUBJECT CLASSIFICATION
47A52; 49N45; 65J20

1. Introduction

In the recent years, considerable progress has been made concerning the analysis of convex Tikhonov regularization in various settings. Existence, stability, and convergence have been treated exhaustively in different settings including that of non-linear problems in Banach spaces with different similarity and regularization terms. Moreover, starting with the paper [1], the questions of reconstruction accuracy and asymptotic error estimates have gradually been answered.

The setting of [1], which we will also pursue in this paper, is that of the stable solution of a linear, but noisy and ill-posed, operator equation

$$Fu = v^\delta$$

by means of Tikhonov regularization

$$u_\alpha^\delta = \arg\min_u \left(\frac{1}{2} \|Fu - v^\delta\|^2 + \alpha \mathcal{R}(u) \right), \tag{1}$$

CONTACT Markus Grasmair markus.grasmair@ntnu.no Department of Mathematical Sciences, Norwegian University of Science and Technology, N-7491 Trondheim, Norway.

with a quadratic similarity term but the convex and lower semi-continuous regularization term $\mathcal{R}$. It was shown in [1] that the source condition

$$\xi^\dagger = F^* \omega^\dagger \in \partial\mathcal{R}(u^\dagger),$$

with $u^\dagger$ being the solution of the noise-free equation, implies the error estimate

$$\mathcal{D}_{\xi^\dagger}(u_\alpha^\delta, u^\dagger) \lesssim \delta$$

for a parameter choice $\alpha \sim \delta$. Here $\mathcal{D}_{\xi^\dagger}$ denotes the Bregman distance for the functional $\mathcal{R}$, which is defined as

$$\mathcal{D}_{\xi^\dagger}(u_\alpha^\delta, u^\dagger) = \mathcal{R}(u_\alpha^\delta) - \mathcal{R}(u^\dagger) - \langle \xi^\dagger, u_\alpha^\delta - u^\dagger \rangle.$$

This result can be seen as a direct generalization of the classical result for quadratic regularization with $\mathcal{R}(u) = \frac{1}{2}\|u\|^2$, where we have the convergence rate

$$\|u_\alpha^\delta - u^\dagger\| \lesssim \delta^{1/2} \qquad \text{if} \qquad u^\dagger = F^* \omega^\dagger,$$

again for the parameter choice $\alpha \sim \delta$. This is due to the fact that the subdifferential of the regularization term consists in this case of the single element $u^\dagger$, and the Bregman distance is simply the squared norm of the difference of the arguments. The classical results, however, are in fact significantly more general, as they can be easily extended to fractional source conditions leading to rates of the form

$$\|u_\alpha^\delta - u^\dagger\| \lesssim \delta^{\frac{2\nu}{2\nu+1}}$$

if the source condition

$$u^\dagger = (F^* F)^\nu \omega^\dagger \tag{2}$$

holds for some $0 < \nu \leq 1$ and the regularization parameter α is chosen appropriately.

In order to generalize these results to non-linear operators F, the paper [2] introduced the idea of variational inequalities, which were later modified in [3,4] in order to deal with lower regularity of the solution as well. As alternative, the idea of approximate source conditions was introduced first for quadratic regularization [5] and then generalized to non-quadratic situations [6]. In their original form, both of these approaches dealt, in the non-quadratic case, only with lower order convergence rates; in the quadratic setting, this would roughly correspond to the classical source condition (2) with $\nu \leq 1/2$. However, modifications were proposed for approximate source conditions in [7,8] and for variational inequalities in [9] in order to accommodate for a higher regularity as well, roughly corresponding to (2) with $1/2 < \nu \leq 1$.

In contrast to the relatively simple source condition (2), variational inequalities and approximate source conditions can be hard to interpret and verify in concrete settings. Thus it would be desirable to obtain restatements in terms of more palpable conditions and to clarify the relation between the different variational and approximate conditions and standard source conditions. For the quadratic case, this relation has been made clear in [10]. For the non-quadratic case, however, such an analysis is, as of now, not available.

1.1. Summary of results

In this article, we will consider convex Tikhonov regularization for linear inverse problems on Hilbert spaces of the form (1). The goal of this article is the derivation of convergence rates, that is, estimates for the difference between the reconstruction u_α^δ and the true solution $u^\dagger$ under the natural generalization

$$\xi^\dagger = (F^*F)^\nu \omega^\dagger \in \partial\mathcal{R}(u^\dagger) \tag{3}$$

of the classical source condition (2) to convex regularization terms. The following theorem briefly summarizes the main results obtained in this paper, see Theorems 7, 10, and 14. For an overview of the notation used here, see Section 2.

Theorem 1. *Assume that a source condition of the form (3) holds for some $0 < \nu \le 1$. Then we have the following convergence rates:*

- *For $0 < \nu \le 1/2$ we have*

$$\mathcal{D}_{\xi^\dagger}(u_\alpha^\delta, u^\dagger) \lesssim \delta^{2\nu} \text{ for } \alpha \sim \delta^{2-2\nu}.$$

- *If $\mathcal{R}$ is p-convex (see Definition 9) and $0 < \nu \le 1/2$ we have*

$$\mathcal{D}_{\xi^\dagger}(u_\alpha^\delta, u^\dagger) \lesssim \delta^{\frac{2\nu p}{p-1+2\nu}} \text{ for } \alpha \sim \delta^{\frac{2p-2-2p\nu+4\nu}{p-1+\nu}}.$$

- *If $\mathcal{R}$ is q-coconvex (see Definition 12) and $1/2 \le \nu \le 1$ we have*

$$\mathcal{D}^{\mathrm{sym}}_{\xi_\alpha^\delta, \xi^\dagger}(u_\alpha^\delta, u^\dagger) \lesssim \delta^{\frac{2\nu q}{1+2\nu q-2\nu}} \text{ for } \alpha \sim \delta^{\frac{2+2\nu q-4\nu}{1+2\nu q-2\nu}}.$$

In the case of quadratic regularization with $\mathcal{R}(u) = \frac{1}{2}\|u\|^2$, all of these results coincide with the classical results found, for instance, in [11]. In Section 6, we will, in addition, discuss the implications for several examples of non-quadratic regularization terms.

2. Mathematical preliminaries

Let U and V be Hilbert spaces and $F : U \to V$ a bounded linear operator. Moreover, let $\mathcal{R} : U \to [0, +\infty]$ be a convex, lower semi-continuous and coercive functional. Given some data $v \in V$, we consider the stable, approximate solution of the equation $Fu = v$ by means of non-quadratic Tikhonov regularization, that is, by minimizing the functional

$$\mathcal{T}_\alpha(u, v) := \frac{1}{2}\|Fu - v\|^2 + \alpha\mathcal{R}(u).$$

More precisely, we assume that $v^\dagger \in V$ is some "true" data, but that we are only given noisy data $v^\delta \in V$ satisfying

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

for some noise level $\delta > 0$. Moreover, we denote the true, that is, $\mathcal{R}$-minimizing, solution of the noise-free equation $Fu = v^\dagger$ by

$$u^\dagger :\in \arg\min_u \{\mathcal{R}(u) : Fu = v^\dagger\}.$$

Our main goal is the estimation of the worst case reconstruction error that can occur for fixed noise level $\delta > 0$ and true data $v^\dagger$. That is, we want to estimate

$$\sup\{D(u_\alpha^\delta, u^\dagger) : v^\delta \in V, u_\alpha^\delta \in \arg\min_u \mathcal{T}_\alpha(u, v^\delta), \|v^\dagger - v^\delta\| \leq \delta\}$$

where $D : U \times U \to [0, +\infty]$ is some distance like measure. In the following results we will mostly use the Bregman distance with respect to the regularization functional $\mathcal{R}$, which is defined as

$$\mathcal{D}_\xi(\tilde{u}, u) := \mathcal{R}(\tilde{u}) - \mathcal{R}(u) - \langle \xi, \tilde{u} - u \rangle,$$

where

$$\xi \in \partial\mathcal{R}(u)$$

is some sub-gradient of $\mathcal{R}$ at u. In addition, we will consider the symmetric Bregman distance

$$\mathcal{D}_{\xi, \tilde{\xi}}^{\mathrm{sym}} := \mathcal{D}_\xi(\tilde{u}, u) + \mathcal{D}_{\tilde{\xi}}(u, \tilde{u}) = \langle \xi - \tilde{\xi}, u - \tilde{u} \rangle$$

for

$$\xi \in \partial\mathcal{R}(u) \text{ and } \tilde{\xi} \in \partial\mathcal{R}(\tilde{u}),$$

as well as the norm in some instances.

2.1. Existence, convergence, and stability

It is well known that Tikhonov regularization with a convex, lower semi-continuous, and coercive regularization term is a well-defined regularization method. That is, the following results hold (see [12, Thms. 3.22, 3.23, 3.26]):

- For every $v \in V$ and every $\alpha>0$, the functional $\mathcal{T}_\alpha(\cdot, v)$ attains its minimum.
- Assume that $v_k \to v \in V$ and $\alpha_k \to \alpha>0$, and let $u_k \in \arg\min_u \mathcal{T}_{\alpha_k}(u, v_k)$. Then the sequence u_k has a weakly convergent sub-sequence. Moreover, if $\bar{u}$ is the weak limit of any weakly convergent sub-sequence $(u_{k'})$, then
$$\bar{u} \in \arg\min_u \mathcal{T}_\alpha(u, v) \qquad \text{and} \qquad \mathcal{R}(u_{k'}) \to \mathcal{R}(\bar{u}).$$
- Assume that
$$\delta_k \to 0, \qquad \alpha_k \to 0, \qquad \text{and} \qquad \delta_k^2/\alpha_k \to 0. \tag{4}$$
Let moreover $v_k \in V$ satisfy $\|v_k - v^\dagger\| \le \delta_k$, and let $u_k \in \arg\min_u \mathcal{T}_{\alpha_k}(u, v_k)$. Then the sequence u_k has a sub-sequence $(u_{k'})$ that converges weakly to some $\mathcal{R}$-minimizing solution $\bar{u}$ of the equation $Fu = v^\dagger$ and $\mathcal{R}(u_{k'}) \to \mathcal{R}(\bar{u})$.

Remark 2. If the functional $\mathcal{T}_\alpha(\cdot, v)$ is strictly convex, which is the case, if and only if the restriction of $\mathcal{R}$ to the kernel of F is strictly convex, then the minimizer of $\mathcal{T}_\alpha(\cdot, v)$ as well as the $\mathcal{R}$-minimizing solution of $Fu = v^\dagger$ are unique. In such a case, a standard sub-sequence argument shows that the whole sequences u_k converge weakly to $\bar{u}$.

Remark 3. The fact that u_α^δ minimizes the Tikhonov functional $\mathcal{T}_\alpha(\cdot, v^\delta)$ implies that

$$\frac{1}{2}\|Fu_\alpha^\delta - v^\delta\|^2 + \alpha\mathcal{R}(u_\alpha^\delta) \le \frac{1}{2}\|Fu^\dagger - v^\delta\|^2 + \alpha\mathcal{R}(u^\dagger) \le \frac{\delta^2}{2} + \alpha\mathcal{R}(u^\dagger), \tag{5}$$

which in turn implies in particular that

$$\mathcal{R}(u_\alpha^\delta) \le \frac{\delta^2}{2\alpha} + \mathcal{R}(u^\dagger).$$

Because of the coercivity of $\mathcal{R}$, it follows that there exists some constant $R = R(\delta^2/\alpha, u^\dagger)$ only depending on the ratio δ^2/α and the true solution $u^\dagger$ (or, rather, the function value $\mathcal{R}(u^\dagger)$ at the true solution) such that

$$\|u_\alpha^\delta\| \le R(\delta^2/\alpha, u^\dagger). \tag{6}$$

We will in the following always be interested in the case where $u^\dagger$ is a fixed $\mathcal{R}$-minimizing solution of $Fu = v^\dagger$ and the noise level δ is small and thus, due to the requirement (4) on the regularization parameter, also the ratio δ^2/α. Therefore, we can always assume that the set of regularized solutions u_α^δ is bounded.

Remark 4. Throughout this paper, we assume that the regularization term $\mathcal{R}$ is coercive, as this guarantees the well-posedness of the regularization method as well as the bound (6), which is needed for the derivation of the convergence rates later on. However, both of these can also be guaranteed under the weaker condition that the Tikhonov functional $\mathcal{T}_\alpha(\cdot, v)$ is coercive for any or, equivalently, every $\alpha>0$ and $v \in V$. For the well-posedness see again [12, Thms. 3.22, 3.23, 3.26]; the bound follows from the inequality (cf. (5))

$$\frac{1}{2}\|Fu_\alpha^\delta - v\|^2 \le \|Fu_\alpha^\delta - v^\delta\|^2 + \|v - v^\delta\|^2 \le \delta^2 + 2\alpha\mathcal{R}(u^\dagger) + \|v - v^\delta\|^2$$

and the fact that $v^\delta \to v^\dagger$ implying that $\|v - v^\delta\|$ remains bounded for every fixed $v \in V$. Thus all the results of this paper remain valid under this more general coercivity condition.

In particular, this holds for regularization with (higher order) homogeneous Sobolev norms or (higher order) total variation

$$\mathcal{R}(u) = \|\nabla^\ell(u)\|_{L^p}^p \text{ or } \mathcal{R}(u) = |D^\ell(u)|(\Omega)$$

with $\ell \in \mathbb{N}$ and $1 < p < +\infty$ provided that the domain Ω is connected and the kernel of F does not contain any polynomials of degree at most $\ell-1$. See for instance [13,14] for the total variation case, [12, Prop. 3.66, 3.70] for quadratic Sobolev and total variation regularization, and [15] for the general, abstract case.

2.2. An interpolation inequality

All of the convergence rate results in this paper are based at some point on the following interpolation inequality, which can, for instance, be found in [16, p. 47]:

Lemma 5. *For all $0 \le \nu \le 1/2$ and all $u \in U$ we have*

$$\|(F^*F)^\nu u\| \le \|Fu\|^{2\nu}\|u\|^{1-2\nu}. \tag{7}$$

More precisely, we will make use of the following result:

Corollary 6. *Let $0 \le \nu \le 1/2$ and assume that $\xi \in U$ satisfies*

$$\xi = (F^*F)^\nu \omega$$

for some $\omega \in U$. Then

$$\langle \xi, u \rangle \leq \|\omega\| \|Fu\|^{2\nu} \|u\|^{1-2\nu} \tag{8}$$

for all $u \in U$.

Proof. With the interpolation inequality (7) we have

$$\langle \xi, u \rangle = \langle (F^*F)^\nu \omega, u \rangle = \langle \omega, (F^*F)^\nu u \rangle \leq \|\omega\| \|(F^*F)^\nu u\| \leq \|\omega\| \|Fu\|^{2\nu} \|u\|^{1-2\nu},$$

which proves the assertion. □

3. Basic convergence rates

We consider first the case of a lower order fractional source condition of the form

$$\xi^\dagger \in \operatorname{Ran}(F^*F)^\nu \cap \partial \mathcal{R}(u^\dagger)$$

with $0 < \nu \leq 1/2$ without any additional conditions on the regularization term $\mathcal{R}$. The limiting case $\nu = 1/2$ can be equivalently written as the more standard source condition $\xi^\dagger \in \operatorname{Ran}(F^*) \cap \partial \mathcal{R}(u^\dagger)$, for which it is well known that one obtains a convergence rate

$$\mathcal{D}_{\xi^\dagger}(u_\alpha^\delta, u^\dagger) \lesssim \delta \qquad \text{for} \qquad \alpha \sim \delta.$$

The following result shows that a weaker source condition leads to a correspondingly slower convergence.

Theorem 7. *Assume that there exists*

$$\xi^\dagger := (F^*F)^\nu \omega^\dagger \in \partial \mathcal{R}(u^\dagger)$$

for some $0 < \nu \leq 1/2$. *Then*

$$\mathcal{D}_{\xi^\dagger}(u_\alpha^\delta, u^\dagger) \lesssim C_1 \frac{\delta^2}{\alpha} + C_2 \delta^{2\nu} + C_3 \alpha^{\frac{\nu}{1-\nu}}.$$

for some constants C_1, C_2, $C_3 > 0$ *whenever* δ^2/α *is bounded. In particular, one obtains with a parameter choice*

$$\alpha(\delta) \sim \delta^{2-2\nu}$$

a convergence rate

$$\mathcal{D}_{\xi^\dagger}(u_\alpha^\delta, u^\dagger) \lesssim \delta^{2\nu}.$$

Proof. We will only consider the case $0 < \nu < 1/2$, the case $\nu = 1/2$ having already been treated in [1].

Since $\xi^\dagger = (F^*F)^\nu \omega^\dagger$, we can apply the interpolation inequality (8), which yields that

$$\langle \xi^\dagger, u^\dagger - u \rangle \le \|\omega^\dagger\| \|F(u^\dagger - u)\|^{2\nu} \|u - u^\dagger\|^{1-2\nu}.$$

Moreover, the fact that u_α^δ minimizes the Tikhonov functional implies that

$$\frac{1}{2}\|Fu_\alpha^\delta - v^\delta\|^2 + \alpha\mathcal{R}(u_\alpha^\delta) \le \frac{1}{2}\|Fu^\dagger - v^\delta\|^2 + \alpha\mathcal{R}(u^\dagger) \le \frac{1}{2}\delta^2 + \alpha\mathcal{R}(u^\dagger).$$

Thus

$$\begin{aligned}\mathcal{D}_{\xi^\dagger}(u_\alpha^\delta; u^\dagger) &= \mathcal{R}(u_\alpha^\delta) - \mathcal{R}(u^\dagger) - \langle \xi^\dagger, u_\alpha^\delta - u^\dagger \rangle \\ &\le \frac{\delta^2}{2\alpha} - \frac{1}{2\alpha}\|Fu_\alpha^\delta - v^\delta\|^2 + \|\omega\| \|F(u^\dagger - u_\alpha^\delta)\|^{2\nu} \|u^\dagger - u_\alpha^\delta\|^{1-2\nu}.\end{aligned} \tag{9}$$

Using Remark 3 we see that the term $\|u^\dagger - u_\alpha^\delta\|$ stays bounded. Using the fact that

$$\|F(u^\dagger - u_\alpha^\delta)\|^{2\nu} \le \|Fu_\alpha^\delta - v^\delta\|^{2\nu} + \delta^{2\nu},$$

we obtain thus from (9) the estimate

$$\mathcal{D}_{\xi^\dagger}(u_\alpha^\delta; u^\dagger) \le \frac{\delta^2}{2\alpha} + C\delta^{2\nu} - \frac{1}{2\alpha}\|Fu_\alpha^\delta - v^\delta\|^2 + C\|Fu_\alpha^\delta - v^\delta\|^{2\nu}$$

for some $C > 0$. Using Young's inequality $ab \le a^p/p + b^{p_*}/p_*$, we see that

$$C\|Fu_\alpha^\delta - v^\delta\|^{2\nu} \le \frac{1}{2\alpha}\|Fu_\alpha^\delta - v^\delta\|^2 + \tilde{C}\alpha^{\frac{\nu}{1-\nu}}$$

for some $\tilde{C} > 0$, and thus

$$\mathcal{D}_{\xi^\dagger}(u_\alpha^\delta; u^\dagger) \le \frac{\delta^2}{2\alpha} + C\delta^{2\nu} + \tilde{C}\alpha^{\frac{\nu}{1-\nu}}.$$

Now the rate follows immediately by inserting the parameter choice $\alpha \sim \delta^{2-2\nu}$. □

Remark 8. In quadratic Tikhonov regularization with

$$\mathcal{R}(u) = \frac{1}{2}\|u\|^2$$

we have that

$$\partial\mathcal{R}(u^\dagger) = u^\dagger \text{ and } \mathcal{D}_{u^\dagger}(u, u^\dagger) = \frac{1}{2}\|u - u^\dagger\|^2.$$

Thus the condition of Theorem 7 reduces to the classical (lower order) source condition

$$u^\dagger \in \operatorname{Ran}(F^*F)^\nu \text{ with } 0 < \nu \le 1/2.$$

The convergence rate obtained in Theorem 7, however, would be

$$\|u_\alpha^\delta - u^\dagger\| \lesssim \delta^\nu \text{ with } \alpha \sim \delta^{2-2\nu}.$$

In contrast, it is well known (see e.g [11]) that a parameter choice

$$\alpha \sim \delta^{\frac{2}{2\nu+1}}$$

leads to a convergence rate

$$\|u_\alpha^\delta - u^\dagger\| \lesssim \delta^{\frac{2\nu}{2\nu+1}}.$$

Since $\nu > 2\nu/(2\nu+1)$ for $0 < \nu < 1/2$, this convergence rate is faster than the one obtained in the Theorem 7. The reason for this discrepancy can be found in the inequality (9), after which we estimate the term $\|u^\dagger - u_\alpha^\delta\|$ simply by a constant. Here better estimates are possible, if we can use some power of the Bregman distance in order to bound this term from above. For quadratic regularization, this is obviously possible, as the Bregman distance is essentially the squared norm. More general instances of this situation will be discussed in the following section.

4. Convergence rates for *p*-convex functionals

As discussed above, in order to obtain stronger results, we need to require a stronger form of convexity for the regularization term $\mathcal{R}$.

Definition 9. Let $1 \le p < +\infty$. We say that the functional $\mathcal{R} : U \to [0, +\infty]$ is locally p-convex, if there exists for each $u \in \operatorname{dom}\partial\mathcal{R}$ and every $R > 0$ some constant $C = C(u, R) > 0$ such that

$$C\|\tilde{u} - u\|^p \le \mathcal{D}_\xi(\tilde{u}, u)$$

for all $\xi \in \partial\mathcal{R}(u)$ and all $\tilde{u} \in U$ with $\|\tilde{u} - u\| \le R$.

Theorem 10. *Assume that $\mathcal{R}$ is locally p-convex for some $p \ge 1$ and that there exists*

$$\xi^\dagger := (F^*F)^\nu \omega^\dagger \in \partial\mathcal{R}(u^\dagger)$$

for some $0 < \nu < 1/2$. Then there exist constants C_1, $C_2 > 0$ such that

$$\mathcal{D}_{\xi^\dagger}(u_\alpha^\delta, u^\dagger) \le C_1 \frac{\delta^2}{\alpha} + C_2 \alpha^{\frac{\nu p}{p-1-p\nu+2\nu}}$$

whenever δ^2/α is bounded. In particular, we obtain with a parameter choice

$$\alpha(\delta) \sim \delta^{\frac{2p-2-2p\nu+4\nu}{p-1+\nu}}$$

the convergence rate

$$\mathcal{D}_{\xi^\dagger}(u_\alpha^\delta, u^\dagger) \lesssim \delta^{\frac{2\nu p}{p-1+2\nu}}.$$

Proof. As in the proof of Theorem 7 we obtain the estimate (cf. inequality (9))

$$\mathcal{D}_{\xi^\dagger}(u_\alpha^\delta; u^\dagger) \leq \frac{\delta^2}{2\alpha} - \frac{1}{2\alpha}\|Fu_\alpha^\delta - v^\delta\|^2 + \|\omega\|\|F(u^\dagger - u_\alpha^\delta)\|^{2\nu}\|u^\dagger - u_\alpha^\delta\|^{1-2\nu}.$$

Again, it follows from Remark 3 that we can assume the term $\|u^\dagger - u_\alpha^\delta\|$ to be bounded. Thus the local p-convexity of $\mathcal{R}$ implies the existence of a constant C such that

$$\|u^\dagger - u_\alpha^\delta\| \leq C\mathcal{D}_{\xi^\dagger}(u_\alpha^\delta, u^\dagger)^{\frac{1}{p}}$$

and we obtain the estimate

$$\mathcal{D}_{\xi^\dagger}(u_\alpha^\delta; u^\dagger) \leq \frac{\delta^2}{2\alpha} - \frac{1}{2\alpha}\|Fu_\alpha^\delta - v^\delta\|^2 + C^{1-2\nu}\|\omega\|\|F(u^\dagger - u_\alpha^\delta)\|^{2\nu}\mathcal{D}_{\xi^\dagger}(u_\alpha^\delta, u^\dagger)^{\frac{1-2\nu}{p}}. \tag{10}$$

We now apply Young's inequality

$$abc \leq \frac{1}{r}a^r + \frac{1}{s}b^s + \frac{1}{t}c^t \text{ for } a,\ b,\ c>0 \text{ and } r,\ s,\ t>1 \text{ with } \frac{1}{r}+\frac{1}{s}+\frac{1}{t} = 1 \tag{11}$$

with

$$\begin{aligned} a &= C^{1-2\nu}\frac{(4\alpha)^\nu\|\omega^\dagger\|}{\nu^\nu}, \quad r = \frac{p}{p-1-p\nu+2\nu}, \\ b &= \frac{\nu^\nu}{(4\alpha)^\nu}\|F(u^\dagger - u_\alpha^\delta)\|^{2\nu}, \quad s = \frac{1}{\nu}, \\ c &= D_{\xi^\dagger}(u_\alpha^\delta, u^\dagger)^{\frac{1-2\nu}{p}}, \quad t = \frac{p}{1-2\nu}, \end{aligned}$$

which results in the bound

$$\|\omega^\dagger\|\|F(u^\dagger - u_\alpha^\delta)\| \leq \tilde{C}\alpha^{\frac{\nu p}{p-1-p\nu+2\nu}} + \frac{1}{4\alpha}\|F(u^\dagger - u_\alpha^\delta)\|^2 + \frac{1-2\nu}{p}\mathcal{D}_{\xi^\dagger}(u_\alpha^\delta, u^\dagger) \tag{12}$$

for some constant $\tilde{C}>0$. Using that

$$\|F(u^\dagger - u_\alpha^\delta)\|^2 \leq 2\|Fu_\alpha^\delta - v^\delta\|^2 + 2\|Fu^\dagger - v^\delta\|^2 \leq 2\|Fu_\alpha^\delta - v^\delta\|^2 + 2\delta^2,$$

and combining (10) with (12), we obtain the required inequality

$$\mathcal{D}_{\xi^\dagger}(u_\alpha^\delta, u^\dagger) \leq C_1\frac{\delta^2}{\alpha} + C_2\alpha^{\frac{\nu p}{p-1-p\nu+2\nu}}$$

for some C_1, $C_2 > 0$. The two terms on the right hand side of this estimate balance for

$$\alpha \sim \delta^{\frac{2p-2-2p\nu+4\nu}{p-1+2\nu}},$$

in which case we obtain the convergence rate

$$\mathcal{D}_{\xi^\dagger}(u_\alpha^\delta, u^\dagger) \lesssim \delta^{\frac{2\nu p}{p-1+2\nu}}.$$

□

Remark 11. Assume that the assumptions of Theorem 10 are satisfied. Because of the local p-convexity of $\mathcal{R}$, we then obtain in addition a convergence rate in terms of the norm of the form

$$\|u_\alpha^\delta - u^\dagger\| \lesssim \delta^{\frac{2\nu}{p-1+2\nu}}.$$

In the particular case of a 2-convex regularization term, we recover the familiar convergence rate

$$\|u_\alpha^\delta - u^\dagger\| \lesssim \delta^{\frac{2\nu}{1+2\nu}} \text{ for } \xi^\dagger \in \operatorname{Ran}(F^*F)^\nu, 0 < \nu \leq 1/2,$$

with a parameter choice

$$\alpha \sim \delta^{\frac{2}{1+2\nu}},$$

which is the same as we obtain for quadratic Tikhonov regularization (cf. Remarks 8).

5. Higher order rates

We will now consider higher order source conditions

$$\xi^\dagger \in \operatorname{Ran}(F^*F)^\nu \cap \partial\mathcal{R}(u^\dagger) \text{ with } \frac{1}{2} < \nu \leq 1.$$

Here it turns out that a strong type of convexity appears not to be needed to obtain higher order convergence rates. Instead, it is the convexity of the conjugate of the regularization term $\mathcal{R}$ that needs to be controlled.

Definition 12. Let $1 \leq q < +\infty$. We say that the functional $\mathcal{R} : U \to [0, +\infty]$ is locally q-coconvex, if there exists for all $R > 0$ some constant $C = C(R) > 0$ such that

$$C\|\xi_1 - \xi_2\|^q \leq \mathcal{D}^{\text{sym}}_{\xi_1,\xi_2}(u_1, u_2) = \langle \xi_1 - \xi_2, u_1 - u_2 \rangle$$

for all u_1, $u_2 \in \operatorname{dom} \partial\mathcal{R}$ with $\|u_i\| \leq R$, where

$$\xi_1 \in \partial\mathcal{R}(u_1) \text{ and } \xi_2 \in \partial\mathcal{R}(u_2).$$

Remark 13. Instead of the original functional $\mathcal{R}$, we can also consider its convex conjugate $\mathcal{R}^*$ and the dual Bregman distances

$$\mathcal{D}_u^*(\tilde{\xi}, \xi) = \mathcal{R}^*(\tilde{\xi}) - \mathcal{R}^*(\xi) - \langle u, \tilde{\xi} - \xi \rangle \text{ with } u \in \partial\mathcal{R}^*(\xi)$$

and

$$\mathcal{D}_{u,\tilde{u}}^{\mathrm{sym},*}(\xi, \tilde{\xi}) := \mathcal{D}_u^*(\tilde{\xi}, \xi) + \mathcal{D}_{\tilde{u}}^*(\xi, \tilde{\xi}) \text{ with } u \in \partial\mathcal{R}^*(\xi) \text{ and } \tilde{u} \in \partial\mathcal{R}^*(\tilde{\xi}).$$

Then we see that the primal and dual symmetric Bregman distances are identical in the sense that

$$\mathcal{D}_{\xi,\tilde{\xi}}^{\mathrm{sym}}(u, \tilde{u}) = \langle \xi - \tilde{\xi}, u - \tilde{u} \rangle = \mathcal{D}_{u,\tilde{u}}^{\mathrm{sym},*}(\xi, \tilde{\xi}).$$

As a consequence, the q-coconvexity of $\mathcal{R}$ is equivalent to the q-convexity of $\mathcal{R}^*$. Also, we note that 2-coconvexity of $\mathcal{R}$ is the same as cocoercivity of the subgradient $\partial\mathcal{R}$ (cf. [17, Sec. 4.2]).

Theorem 14. *Assume that $\mathcal{R}$ is locally q-coconvex for some $q \geq 1$ and that*

$$\xi^\dagger := (F^*F)^\nu \eta^\dagger \in \partial\mathcal{R}(u^\dagger)$$

for some $1/2 < \nu \leq 1$. Then there exist constants C_1, $C_2 > 0$ such that

$$\mathcal{D}_{\xi_\alpha^\delta, \xi^\dagger}^{\mathrm{sym}}(u_\alpha^\delta, u^\dagger) \leq C_1 \frac{\delta^2}{\alpha} + C_2 \alpha^{\frac{q\nu}{1+\nu q - 2\nu}}. \tag{13}$$

whenever δ^2/α is bounded. In particular, we obtain with a parameter choice

$$\alpha \sim \delta^{\frac{2+2\nu q - 4\nu}{1+2\nu q - 2\nu}}$$

the convergence rate

$$\mathcal{D}_{\xi_\alpha^\delta, \xi^\dagger}^{\mathrm{sym}}(u_\alpha^\delta, u^\dagger) \lesssim \delta^{\frac{2\nu q}{1+2\nu q - 2\nu}}.$$

Proof. Denote

$$\mu = \nu - \frac{1}{2}.$$

Since

$$\mathrm{Ran}(F^*F)^\nu = \mathrm{Ran}(F^*F)^{\mu + \frac{1}{2}} = \mathrm{Ran}(F^*(FF^*)^\mu),$$

it follows that we can write

$$\xi^\dagger = F^*\omega^\dagger \text{ with } \omega^\dagger = (FF^*)^\mu \tilde{\eta}^\dagger$$

for some $\tilde{\eta}^\dagger \in U$.

Because u_α^δ is a minimizer of the Tikhonov functional $\mathcal{T}_\alpha(\cdot, v^\delta)$, it satisfies the first order optimality condition

$$F^*(Fu_\alpha^\delta - v^\delta) + \alpha\partial\mathcal{R}(u_\alpha^\delta) \ni 0.$$

Denoting by

$$\xi_\alpha^\delta \in \partial\mathcal{R}(u_\alpha^\delta)$$

the corresponding subgradient of $\mathcal{R}$, it follows that

$$-\alpha\xi_\alpha^\delta = F^*(Fu_\alpha^\delta - v^\delta).$$

Or, we can write

$$\xi_\alpha^\delta = F^*\omega_\alpha^\delta \text{ with } -\alpha\omega_\alpha^\delta = Fu_\alpha^\delta - v^\delta. \tag{14}$$

As a consequence, we have

$$\begin{aligned}
\mathcal{D}^{\mathrm{sym}}_{\xi_\alpha^\delta,\xi^\dagger}(u_\alpha^\delta, u^\dagger) &= \langle \xi_\alpha^\delta - \xi^\dagger, u_\alpha^\delta - u^\dagger\rangle \\
&= \langle F^*\omega_\alpha^\delta - F^*\omega^\dagger, u_\alpha^\delta - u^\dagger\rangle \\
&= \langle \omega_\alpha^\delta - \omega^\dagger, Fu_\alpha^\delta - Fu^\dagger\rangle \\
&= \langle \omega_\alpha^\delta - \omega^\dagger, Fu_\alpha^\delta - v^\delta\rangle + \langle \omega_\alpha^\delta - \omega^\dagger, v^\delta - v^\dagger\rangle \\
&= -\alpha\langle \omega_\alpha^\delta - \omega^\dagger, \omega_\alpha^\delta\rangle + \langle \omega_\alpha^\delta - \omega^\dagger, v^\delta - v^\dagger\rangle \\
&= -\alpha\|\omega_\alpha^\delta - \omega^\dagger\|^2 - \alpha\langle \omega_\alpha^\delta - \omega^\dagger, \omega^\dagger\rangle + \langle \omega_\alpha^\delta - \omega^\dagger, v^\delta - v^\dagger\rangle \\
&\leq -\alpha\|\omega_\alpha^\delta - \omega^\dagger\|^2 - \alpha\langle \omega_\alpha^\delta - \omega^\dagger, \omega^\dagger\rangle + \delta\|\omega_\alpha^\delta - \omega^\dagger\|.
\end{aligned} \tag{15}$$

We next use the interpolation inequality and the definitions of ω_α^δ and $\omega^\dagger$ and obtain

$$\begin{aligned}
-\langle \omega_\alpha^\delta - \omega^\dagger, \omega^\dagger\rangle &= -\langle \omega_\alpha^\delta - \omega^\dagger, (FF^*)^\mu \tilde{\eta}^\dagger\rangle \\
&\leq \|\tilde{\eta}^\dagger\|\|F^*(\omega_\alpha^\delta - \omega^\dagger)\|^{2\mu}\|\omega_\alpha^\delta - \omega^\dagger\|^{1-2\mu} \\
&= \|\tilde{\eta}^\dagger\|\|\xi_\alpha^\delta - \xi^\dagger\|^{2\mu}\|\omega_\alpha^\delta - \omega^\dagger\|^{1-2\mu}.
\end{aligned} \tag{16}$$

Now we can use the local q-coconvexity of $\mathcal{R}$ and the boundedness of u_α^δ (see Remark 3) to estimate

$$\|\xi_\alpha^\delta - \xi^\dagger\| \leq C\mathcal{D}^{\mathrm{sym}}_{\xi_\alpha^\delta,\xi^\dagger}(u_\alpha^\delta, u^\dagger)^{1/q}$$

and obtain from (15) and (16) the bound

$$\begin{aligned}
&\mathcal{D}^{\mathrm{sym}}_{\xi_\alpha^\delta,\xi^\dagger}(u_\alpha^\delta, u^\dagger) \\
&\leq C\alpha\|\tilde{\eta}^\dagger\|\mathcal{D}^{\mathrm{sym}}_{\xi_\alpha^\delta,\xi^\dagger}(u_\alpha^\delta, u^\dagger)^{\frac{2\mu}{q}}\|\omega_\alpha^\delta - \omega^\dagger\|^{1-2\mu} + \delta\|\omega_\alpha^\delta - \omega^\dagger\| - \alpha\|\omega_\alpha^\delta - \omega^\dagger\|^2.
\end{aligned} \tag{17}$$

In the following, we will only treat the more difficult case $\mu < 1/2$. For $\mu = 1/2$, the argumentation is similar but simpler, due to the absence of the term $\|\omega_\alpha^\delta - \omega^\dagger\|^{1-2\mu}$ in the first product on the right hand side of (17).

We use first the inequality

$$\delta\|\omega_\alpha^\delta - \omega^\dagger\| \le \frac{\delta^2}{2\alpha} + \frac{\alpha}{2}\|\omega_\alpha^\delta - \omega^\dagger\|^2$$

and then the three term Young inequality (11) with

$$\begin{aligned} a &= C(1-2\mu)^{\frac{1-2\mu}{2}}\|\tilde{\eta}^\dagger\|\alpha^{\frac{1+2\mu}{2}}, \quad r = \frac{2q}{q+2\mu q-4\mu}, \\ b &= \frac{\alpha^{\frac{1-2\mu}{2}}}{(1-2\mu)^{\frac{1-2\mu}{2}}}\|\omega_\alpha^\delta - \omega^\dagger\|^{1-2\mu}, \quad s = \frac{2}{1-2\mu}, \\ c &= \mathcal{D}^{\mathrm{sym}}_{\xi_\alpha^\delta,\xi^\dagger}(u_\alpha^\delta, u^\dagger)^{\frac{2\mu}{q}}, \quad t = \frac{q}{2\mu}. \end{aligned}$$

Then we obtain from (17) that

$$\mathcal{D}^{\mathrm{sym}}_{\xi_\alpha^\delta,\xi^\dagger}(u_\alpha^\delta, u^\dagger) \le C_1\frac{\delta^2}{\alpha} + C_2\alpha^{\frac{q+2\mu q}{q+2\mu q-4\mu}}.$$

Again, balancing the two terms on the right hand side leads to a parameter choice

$$\alpha \sim \delta^{\frac{q+2\mu q-4\mu}{q+2\mu q-2\mu}}$$

and a convergence rate

$$\mathcal{D}^{\mathrm{sym}}_{\xi_\alpha^\delta,\xi^\dagger}(u_\alpha^\delta, u^\dagger) \lesssim \delta^{\frac{q+2\mu q}{q+2\mu q-2\mu}}.$$

Replacing again μ by $\nu-\frac{1}{2}$, we obtain the results claimed in the statement of the theorem. □

Remark 15. The equations (14) are just the KKT conditions for the optimization problem $\min_u \mathcal{T}_\alpha(u, v^\delta)$, and ω_α^δ can be just seen as the dual solution of this problem. See also [9], where the connection to a dual Tikhonov functional is discussed.

Remark 16. The lower order rates obtained in Section 4 at first glance appear to be different from the higher order rates of the previous theorem. By reformulating the rates not in terms of q but rather its Hölder conjugate, however, it is possible to use the same formula in both cases. Indeed, assume that $\mathcal{R}$ is locally q-coconvex and that $p = q/(q-1)$ is the Hölder conjugate of q. Then we can write the estimate (13) as

$$\mathcal{D}^{\mathrm{sym}}_{\xi_\alpha^\delta,\xi^\dagger}(u_\alpha^\delta, u^\dagger) \le C_1\frac{\delta^2}{\alpha} + C_2\alpha^{\frac{p\nu}{p-1-p\nu+2\nu}},$$

which is of the same form as the estimate we have obtained in Theorem 10 for p-convex regularization terms and $\nu < 1/2$.

Remark 17. In the case where the regularization term $\mathcal{R}$ is 2-coconvex, the parameter choice and convergence rate simplify to

$$\mathcal{D}^{\mathrm{sym}}_{\xi_\alpha^\delta,\xi^\dagger}(u_\alpha^\delta,u^\dagger)\lesssim\delta^{\frac{4\nu}{1+2\nu}}\qquad\text{for}\qquad\alpha\sim\delta^{\frac{2}{1+2\nu}}.$$

In the case of quadratic Tikhonov regularization, these rates turn out to be identical to the classical rates. Indeed, the quadratic norm is obviously 2-coconvex, since we have for

$$\mathcal{R}(u)=\frac{1}{2}\|u\|^2$$

that

$$\partial\mathcal{R}(u)=\{u\}\ \text{and}\ \mathcal{D}^{\mathrm{sym}}_{u_1,u_2}(u_1,u_2)=\|u_1-u_2\|^2.$$

Moreover, the source condition simply reads as

$$u^\dagger=(F^*F)^\nu\eta^\dagger.$$

Together with Remark 11, which deals with the lower order case, we thus recover the classical result that the source condition

$$u^\dagger\in\mathrm{Ran}(F^*F)^\nu\ \text{for some}\ 0<\nu\leq 1$$

implies the convergence rate

$$\|u_\alpha^\delta-u^\dagger\|\lesssim\delta^{\frac{2\nu}{2\nu+1}}\ \text{with}\ \alpha\sim\delta^{\frac{2}{2\nu+1}}$$

for quadratic Tikhonov regularization.

6. Examples

We now study the implications for four different non-quadratic regularization terms, all with different convexity properties.

6.1. ℓ^p-regularization

We consider first the case where $U=\ell^2(I)$ for some countable index set I, and

$$\mathcal{R}(u)=\frac{1}{p}\|u\|_{\ell^p}^p=\frac{1}{p}\sum_{i\in I}|u_i|^p$$

for some $1<p<2$. Because of the embedding $\ell^p\to\ell^2$ for $p<2$, this term is coercive and thus Tikhonov regularization is well-posed. Also, this regularization term is 2-convex and its conjugate

$$\mathcal{R}^*(\xi) = \frac{1}{p_*}\|\xi\|_{\ell^{p_*}}^{p_*}$$

is p_*-convex with $p_* = p/(p-1)$ being the Hölder conjugate of p, implying that $\mathcal{R}$ is p_*-coconvex (see [18] for all of these results). Moreover,

$$\partial\mathcal{R}(u) = (u_i|u_i|^{p-2})_{i\in I}$$

whenever $u \in \operatorname{dom}\partial\mathcal{R} = \ell^{2(p-1)}$.

Thus the preceding results imply that a source condition

$$\xi^\dagger = (u_i^\dagger|u_i^\dagger|^{p-2})_{i\in I} \in \operatorname{Ran}(F^*F)^\nu$$

leads to a convergence rate

$$\|u_\alpha^\delta - u^\dagger\| \lesssim \delta^{\frac{2\nu}{1+2\nu}} \text{ with } \alpha \sim \delta^{\frac{2}{1+2\nu}} \qquad \text{if } 0 < \nu \leq \frac{1}{2},$$

and

$$\|u_\alpha^\delta - u^\dagger\| \lesssim \delta^{\frac{p\nu}{p-1+2\nu}} \text{ with } \alpha \sim \delta^{\frac{2p-2-2\nu p+4\nu}{p-1+2\nu}} \qquad \text{if } \frac{1}{2} < \nu \leq 1.$$

6.2. L^p-regularization

Next we study the situation where Ω is some bounded domain, $U = L^2(\Omega)$, and

$$\mathcal{R}(u) = \frac{1}{p}\int_\Omega |u(x)|^p \, dx = \frac{1}{p}\|u\|_{L^p}^p$$

for some $2 < p < +\infty$. Here we are in the opposite situation to ℓ^p-regularization in that the exponent has to be larger than 2 for the regularization method to be well-posed. See [19,20] for an application of this type of regularization, albeit completely in a Banach space setting.

In this case the regularization term itself is p-convex, but its conjugate

$$\mathcal{R}^*(u) = \frac{1}{p_*}\|u\|_{L^{p_*}}^{p_*}$$

is 2-convex (see again [18]). Also, we have again the representation of the subgradient of $\mathcal{R}$ as

$$\partial\mathcal{R}(u) = u|u|^{p-2}$$

whenever $u \in \operatorname{dom}\partial\mathcal{R} = L^{2(p-1)}$.

As a consequence, due to the p-convexity and 2-coconvexity of the regularization term, the results above imply that the source condition

$$\xi^\dagger := u|u|^{p-2} \in \operatorname{Ran}(F^*F)^\nu$$

results in the convergence rates

$$\|u_\alpha^\delta - u^\dagger\| \lesssim \delta^{\frac{2\nu}{p-1+2\nu}} \text{ with } \alpha \sim \delta^{\frac{2p-2-2p\nu+4\nu}{p-1+\nu}} \qquad \text{if } 0 < \nu \leq \frac{1}{2},$$

and

$$\|u_\alpha^\delta - u^\dagger\| \lesssim \delta^{\frac{4\nu}{p+2\nu p}} \text{ with } \alpha \sim \delta^{\frac{2}{1+2\nu}} \qquad \text{if } \frac{1}{2} < \nu \leq 1.$$

6.3. Total variation regularization

The next example we consider is total variation regularization with $U = L^2(\Omega), \Omega \subset \mathbb{R}^2$ bounded with Lipschitz boundary, and

$$\mathcal{R}(u) = |Du|(\Omega).$$

As discussed in Remark 4, we have to assume in this case in addition that constant functions are not contained in the kernel of F in order for the regularization method to be well-posed.

In the case of total variation regularization, the regularization term is not strictly convex, which implies that the Bregman distance $\mathcal{D}_\xi(\tilde{u}, u)$ may be zero for $\tilde{u} \neq u$. As a consequence, we cannot bound the Bregman distance from below by any power of the norm, and therefore the total variation is not p-convex for any p. On the other hand, the subdifferentials of $\mathcal{R}$ are in general not single-valued, which implies that the total variation is neither q-coconvex for any q. We thus end up with only the basic results

$$\mathcal{D}_{\xi^\dagger}(u_\alpha^\delta, u^\dagger) \lesssim \delta^{2\nu} \text{ with } \alpha \sim \delta^{2-2\nu}$$

for a source condition

$$\xi^\dagger \in \partial\mathcal{R}(u^\dagger) \cap \operatorname{Ran}(F^*F)^\nu \text{ with } 0 < \nu \leq \frac{1}{2}.$$

6.4. Huber regularization

As final example, we get back to the case $U = \ell^2(I)$ for some countable index set I, but consider now the Huber regularization term

$$\mathcal{R}(u) = \sum_{i \in I} \phi(u_i)$$

with

$$\phi(t) = \begin{cases} \frac{1}{2}t^2 & \text{if } |t| \leq 1, \\ |t| - \frac{1}{2} & \text{if } |t| \geq 1. \end{cases}$$

Because ϕ is not strictly convex, neither is $\mathcal{R}$, and thus $\mathcal{R}$ is not p-convex for any p. However,

$$\mathcal{R}^*(\xi) = \sum_{i\in I} \phi^*(\xi_i)$$

with

$$\phi^*(\zeta) = \begin{cases} \frac{1}{2}\zeta^2 & \text{if } |\zeta| \le 1, \\ +\infty & \text{if } |\zeta > 1, \end{cases}$$

which is obviously 2-convex. Thus the Huber regularization term is not p-convex for any p, but is 2-coconvex.

Moreover, we have that

$$\partial\mathcal{R}(u) = (\rho(u_i))_{i\in I}$$

with

$$\rho(t) = \begin{cases} 1 & \text{if } t \ge 1, \\ t & \text{if } |t| \le 1, \\ -1 & \text{if } t \le -1. \end{cases}$$

Thus we obtain the convergence rates

$$\mathcal{D}_{\xi^\dagger}(u_\alpha^\delta, u^\dagger) \lesssim \delta^{2\nu} \text{ with } \alpha \sim \delta^{2-2\nu} \qquad \text{if } 0 < \nu \le \frac{1}{2}$$

and

$$\mathcal{D}^{\text{sym}}_{\xi_\alpha^\delta, \xi^\dagger}(u_\alpha^\delta, u^\dagger) \lesssim \delta^{\frac{2\nu}{1+2\nu}} \text{ with } \alpha \sim \delta^{\frac{2}{1+2\nu}} \qquad \text{if } \frac{1}{2} \le \nu \le 1,$$

provided that a source condition

$$(\rho(u_i^\dagger))_{i\in I} \in \operatorname{Ran}(F^*F)^\nu$$

is satisfied.

7. Conclusion

In this paper we have studied the implications of classical source conditions of power type to accuracy estimates and convergence rates for non-quadratic Tikhonov regularization. We have seen that very basic results can be easily obtained without any additional conditions concerning, for instance, strong convexity or smoothness of the regularization term. However, these results are not optimal in cases where such additional conditions hold, and they also fail to reproduce the classical results for quadratic regularization methods.

In order to be able to obtain stronger results, we considered the situation where either the regularization term or its convex conjugate is p-convex. In these cases, it is possible to obtain sharper estimates in the low regularity and high regularity regions, respectively. As for quadratic regularization, the split between these two regions occurs at the standard source condition $F^*\omega^\dagger \in \partial\mathcal{R}(u^\dagger)$. In the lower order case, the convexity properties of the primal regularization term $\mathcal{R}$ determine the convergence rates. In the higher order case, however, the proof of the rates relies to some extent on duality and it is therefore the convexity of the dual regularization term $\mathcal{R}^*$ that becomes important. Still, these results match those classically obtained for quadratic regularization, although the approach we have followed here differs significantly from the classical ones.

However, quite a few questions remain open. First, all the results we have discussed here were obtained only for the case of linear inverse problems. It seems reasonable, though, to expect that a refinement of the approach chosen in this paper might lead to convergence rates for non-linear problems as well. This would be particularly desirable for the case of enhanced convergence rates in the high regularity region, where up to now no easily interpretable results are available.

Next, it is well known (see [4, 21]) that sparsity assumptions lead to improved convergence rates of, for instance, order $\delta^{1/p}$ in the case of ℓ^p-regularization with $1 < p < 2$. Therefore, it would make sense to investigate whether sparsity might in general alter and improve error estimates and convergence rates in the case of Hölder type fractional source conditions. For the setting of ℓ^1-regularization, it is known that lower order fractional source conditions $\partial\mathcal{R}(u^\dagger) \cap \operatorname{Ran}(F^*F)^\nu \neq \emptyset$ for any $0 < \nu \leq 1/2$ imply linear convergence rates in the presence of sparsity (see [22]); for ℓ^p-regularization, the effect of such source conditions is still an open problem.

Finally, all these results apply strictly to Hilbert spaces only, as they make use of fractional powers of the operator F and of an interpolation inequality. Since non-quadratic regularization methods become more important in settings without a Hilbert space structure, a generalization of such source conditions together with corresponding convergence rates to Banach spaces would be desirable. All of these points will be subject of further investigation in the future.

Acknowledgements

I want to thank the anonymous referees for their valuable remarks, in particular concerning the relation between the lower and higher order rates.

ORCID

Markus Grasmair http://orcid.org/0000-0002-6116-5584

References

[1] Burger, M., Osher, S. (2004). Convergence rates of convex variational regularization. *Inverse Prob*. 20(5):1411–1421. DOI: 10.1088/0266-5611/20/5/005.

[2] Hofmann, B., Kaltenbacher, B., Pöschl, C., Scherzer, O. (2007). A convergence rates result for Tikhonov regularization in Banach spaces with non-smooth operators. *Inverse Prob*. 23(3):987–1010. DOI: 10.1088/0266-5611/23/3/009.

[3] Boţ, R. I., Hofmann, B. (2010). An extension of the variational inequality approach for obtaining convergence rates in regularization of nonlinear ill-posed problems. *J. Integral Equations Appl*. 22(3):369–392. DOI: 10.1216/JIE-2010-22-3-369.

[4] Grasmair, M. (2010). Generalized Bregman distances and convergence rates for non-convex regularization methods. *Inverse Prob*. 26(11):115014–115016. DOI: 10.1088/0266-5611/26/11/115014.

[5] Hofmann, B. (2006). Approximate source conditions in Tikhonov-Phillips regularization and consequences for inverse problems with multiplication operators. *Math. Meth. Appl. Sci*. 29(3):351–371. DOI: 10.1002/mma.686.

[6] Hein, T. (2008). Convergence rates for regularization of ill-posed problems in Banach spaces by approximate source conditions. *Inverse Prob*. 24(4): 045007–045010. DOI: 10.1088/0266-5611/24/4/045007.

[7] Hein, T. (2009). Tikhonov regularization in Banach spaces—improved convergence rates results. *Inverse Prob*. 25(3):035002. DOI: 10.1088/0266-5611/25/3/035002.

[8] Neubauer, A., Hein, T., Hofmann, B., Kindermann, S., Tautenhahn, U. (2010). Improved and extended results for enhanced convergence rates of Tikhonov regularization in Banach spaces. *Appl. Anal*. 89(11):1729–1743. DOI: 10.1080/00036810903517597.

[9] Grasmair, M. (2013). Variational inequalities and higher order convergence rates for Tikhonov regularisation on Banach spaces. *J. Inverse Ill-Posed Probl*. 21(3):379–394.

[10] Andreev, R., Elbau, P., de Hoop, M. V., Qiu, L., Scherzer, O. (2015). Generalized convergence rates results for linear inverse problems in Hilbert spaces. *Numer. Funct. Anal. Optim*. 36(5):549–566. DOI: 10.1080/01630563.2015.1021422.

[11] Groetsch, C. W. (1984). *The* Theory *of Tikhonov* Regularization *for Fredholm* Equations *of the first kind*. Volume 105 of Research Notes in Mathematics. Boston, MA: Pitman (Advanced Publishing Program).

[12] Scherzer, O., Grasmair, M., Grossauer, H., Haltmeier, M., Lenzen, F. (2009). *Variational Methods in Imaging. Volume 167 of Applied Mathematical Sciences*. New York: Springer.

[13] Acar, R., Vogel, C. R. (1994). Analysis of bounded variation penalty methods for ill-posed problems. *Inverse Prob*. 10(6):1217–1229. DOI: 10.1088/0266-5611/10/6/003.

[14] Vese, L. (2001). A study in the BV space of a denoising-deblurring variational problem. *Appl. Math. Optim*. 44(2):131–161. DOI: 10.1007/s00245-001-0017-7.

[15] Grasmair, M. (2011). Linear convergence rates for Tikhonov regularization with positively homogeneous functionals. *Inverse Prob*. 27(7):075014–075016. DOI: 10.1088/0266-5611/27/7/075014.

[16] Engl, H. W., Hanke, M., Neubauer, A. (1996). *Regularization of Inverse Problems, Volume 375 of Mathematics and Its Applications*. Dordrecht: Kluwer Academic Publishers Group.

[17] Bauschke, H. H., Combettes, P. L. (2011). *Convex Analysis and* Monotone Operator Theory *in Hilbert spaces*. CMS Books in Mathematics/Ouvrages de Mathématiques de la SMC. New York: Springer, With a foreword by Hédy Attouch.

[18] Bonesky, T., Kazimierski, K. S., Maass, P., Schöpfer, F., Schuster, T. (2008). Minimization of Tikhonov functionals in Banach spaces. *Abstr. Appl. Anal.* 2008: 1–19. pages Art. ID 192679, 19, DOI: 10.1155/2008/192679.

[19] Lechleiter, A., Kazimierski, K. S., Karamehmedović, M. (2013). Tikhonov regularization in *Lp* applied to inverse medium scattering. *Inverse Prob.* 29(7):075003–075019. DOI: 10.1088/0266-5611/29/7/075003.

[20] Lechleiter, A., Center for Industrial Mathematics, University of Bremen, 28359 Bremen, Germany, Rennoch, A. (2017). Non-linear Tikhonov regularization in Banach spaces for inverse scattering from anisotropic penetrable media. *Inverse Probl. Imaging*. 11(1):151–176. DOI: 10.3934/ipi.2017008.

[21] Grasmair, M., Haltmeier, M., Scherzer, O. (2008). Sparse regularization with *lq* penalty term. *Inverse Prob.* 24(5):055020–055013. DOI: 10.1088/0266-5611/24/5/055020.

[22] Frick, K., Grasmair, M. (2012). Regularization of linear ill-posed problems by the augmented Lagrangian method and variational inequalities. *Inverse Prob.* 28(10): 104005–104016. DOI: 10.1088/0266-5611/28/10/104005.