Inverse Problems **27** (2011) 025006 (18pp) doi:10.1088/0266-5611/27/2/025006

Sharp converse results for the regularization error using distance functions

Jens Flemming[1], Bernd Hofmann[1] and Peter Mathé[2]

[1] Department of Mathematics, Chemnitz University of Technology, 09107 Chemnitz, Germany
[2] Weierstraß Institute for Applied Analysis and Stochastics, Mohrenstraße 39, 10117 Berlin, Germany

E-mail: jens.flemming@mathematik.tu-chemnitz.de, hofmannb@mathematik.tu-chemnitz.de and mathe@wias-berlin.de

Received 7 September 2010, in final form 15 November 2010
Published 11 January 2011
Online at stacks.iop.org/IP/27/025006

Abstract

In the analysis of ill-posed inverse problems the impact of solution smoothness on accuracy and convergence rates plays an important role. For linear ill-posed operator equations in Hilbert spaces and with focus on the linear regularization schema we will establish relations between different kinds of measuring solution smoothness in a point-wise or integral manner. In particular, we discuss the interplay of distribution functions, profile functions that express the regularization error, index functions generating source conditions and distance functions associated with benchmark source conditions. We show that typically the distance functions and the profile functions carry the same information as the distribution functions, and that this is not the case for general source conditions. The theoretical findings are accompanied with examples exhibiting applications and limitations of the approach. A detailed understanding of solution smoothness will also be helpful for the treatment and convergence analysis of nonlinear ill-posed problems.

1. Introduction

The stable approximate solution of ill-posed inverse problems that can be formulated as linear operator equations

$$Ax = y \tag{1}$$

with an injective and bounded linear operator $A : X \to Y$ mapping between Hilbert spaces X and Y and possessing a non-closed range $\mathcal{R}(A)$ requires regularization, since under the above assumptions the (formal) solution mapping $y \mapsto x = A^{-1}y$ exists for each $y \in \mathcal{R}(A)$, however, this dependence is discontinuous. Precisely, this is the *Moore–Penrose inverse* $A^{\dagger}$, which can be defined in a more general context as a densely defined (unbounded) operator.

Therefore, the solution theory aims at replacing the discontinuous mapping $A^\dagger$ by a family of continuous (bounded) regularization operators $R_\alpha : Y \to X$ indexed by the regularization parameter $\alpha > 0$. This approach is common since the early days of the theory of ill-posed equations, and a seminal treatise along these lines is [6]. The goal is to design families R_α with the property that $R_\alpha y \to A^\dagger y$ as $\alpha \to 0$ whenever $y \in \mathcal{R}(A)$.

Regularization error. The deviation of $R_\alpha y$ from $A^\dagger y$ is called *regularization error*, and we have for $y = Ax$ that

$$f_x(\alpha) := \|A^\dagger y - R_\alpha y\| = \|A^\dagger Ax - R_\alpha Ax\| = \|x - R_\alpha Ax\|, \qquad \alpha > 0.$$

In most cases the family R_α is given by a family of piecewise continuous real functions g_α. By noting that $A^\dagger = (A^*A)^\dagger A^*$ we assign

$$R_\alpha := g_\alpha(A^*A)A^*, \tag{2}$$

where spectral calculus allows us to extend the real functions to operator valued ones. In terms of the family g_α the regularization error reads as $f_x(\alpha) = \|x - g_\alpha(A^*A)A^*Ax\|$, and this function was called *profile function* in [13].

A crucial observation in this context is that the decay rate of $f_x(\alpha)$ as $\alpha \to 0$ depends on smoothness properties of the solution element x. The quantitative relation between smoothness properties of x, given in terms of spectral information, called distribution function below, and the decay of the profile function for a given regularization was first emphasized in [24, 25]. Here we take a more general point of view, and we shall subsume such properties as different kinds of *solution smoothness* of x with respect to the operator $H := A^*A$.

In this study we focus on the understanding of profile functions. It will be clear from the discussion at the end of section 4 that the knowledge of the profile function completely determines the worst-case behavior of the reconstruction error from noisy data.

Distribution function. The *pointwise characteristics* of the solution smoothness of x with respect to the spectrum of H, which contains the complete spectral information of the element, exploits the (right-continuous version of the) spectral *distribution function:*

$$F_x^2(t) := \|E_t x\|^2 := \langle \chi_{(0,t]}(H)\,x, x\rangle = \|\chi_{(0,t]}(H)x\|^2, \qquad 0 < t < \infty, \tag{3}$$

or the equivalent representation as $F_x^2(t) = \int_0^t \mathrm{d}\|E_s x\|^2$ for $t > 0$. Above, we let $\chi_{(0,t]}$ be the characteristic function of the interval $(0, t]$, and $E_t = E_t(H)$, $0 \leqslant t \leqslant \|H\|$, be the spectral resolution of the operator H, i.e. for any (bounded measurable) real function h we have that

$$\|h(H)x\|^2 = \int_0^{\|H\|} h^2(t)\,\mathrm{d}\|E_t x\|^2,$$

we refer to [26, chapter 12] for details on spectral theory of bounded self-adjoint operators in Hilbert space. We should note here that by definition the function F_x is non-decreasing and right-continuous with $\lim_{t\to 0} F_x(t) = 0$. The latter property is a consequence of zero being an accumulation point of the spectrum of H. The distribution function may have jumps at the points of the spectrum of H. In particular, for compact operators H it will be piecewise constant. Moreover, an immediate and well-known observation is the following fact.

Fact 1 ([25, proposition 2.3]). *Given $x \in X$, the increase $F_x(\alpha) = \mathcal{O}(\alpha^\kappa)$ as $\alpha \to 0$, for some $0 < \kappa < 1$, implies that $x \in \mathcal{R}(H^\nu)$ for every $0 < \nu < \kappa$.*

General smoothness. The smoothness of solutions $x \in X$ for ill-posed equations can be expressed mathematically in different ways. The most traditional form is characterized by *source conditions*. In its general version it is assumed that there is a non-decreasing continuous function $\psi : (0, \|H\|] \to \mathbb{R}^+ = (0, \infty)$ with $\lim_{t\to 0} \psi(t) = 0$ that we will call the *index function*, for which

$$x = \psi(H)\, v, \tag{4}$$

with some *source element* $v \in X$. Evidently, (4) can be rewritten as $x \in \mathcal{R}(\psi(H))$. Recent results assert that for every $x \in X$ there is an index function ψ and a source element $v \in X$ satisfying (4), see [15, 22].

Although the concept of source conditions (4) proved to be useful in the error analysis of ill-posed operator equations, the pointwise characteristics (3) contain more precise information of the smoothness of x with respect to the operator H. We exhibit this in case of Tikhonov regularization, where the family R_α is given for $y = Ax$ by $R_\alpha y = (H + \alpha I)^{-1} Hx$, and hence the corresponding profile function is

$$f_x(\alpha) := \alpha \| (H + \alpha I)^{-1} Hx \|, \qquad 0 < \alpha < \infty, \tag{5}$$

which is an index function, well defined and increasing for all $\alpha > 0$. A celebrated converse result establishes a one-to-one correspondence between the distribution and the profile functions in moderate cases.

Fact 2 ([25, theorem 2.1]). *Given $x \in H$, we have that $f_x(\alpha) = \mathcal{O}(\alpha^\kappa)$, for some $0 < \kappa < 1$, if and only if $F_x(\alpha) = \mathcal{O}(\alpha^\kappa)$ as $\alpha \to 0$.*

The key to the proof of fact 2 is the following result. It uses the family $w_\alpha \in X$ of elements given as $w_\alpha := \chi_{(\alpha,\|H\|]}(H)H^{-1}x,\ \alpha > 0$, which will prove useful later, see (A.2) for the general construction.

Lemma 1 ([25], or the original study [7]) **.** *Let w_α be defined as above. The behavior $F_x(\alpha) = \mathcal{O}(\alpha^\kappa)$, for some $0 < \kappa < 1$, yields $\alpha\|w_\alpha\| = \mathcal{O}(\alpha^\kappa)$ as $\alpha \to 0$.*

In contrast, by using power-type source conditions we only have that $f_x(\alpha) = \mathcal{O}(\alpha^\nu)$ provided that $x \in \mathcal{R}(H^\nu)$, again whenever $0 < \nu < 1$. This of course is less accurate than the assertion in fact 2, and it is mentioned in [25] that there are $x \in \mathcal{R}(H^\nu)$ for which $f_x(\alpha) = o(\alpha^\nu)$. We will return to this discussion in corollary 2.

Distance function. In the past years, with the study [10] as well as the subsequent studies in [3, 4, 11, 13, 17] and [9], the lack of information occurring when general source conditions are used, was circumvented by using *distance functions*:

$$d_\psi(R) := \inf\{\|x - \psi(H)v\| : \ \|v\| \leqslant R\}, \qquad 0 \leqslant R < \infty. \tag{6}$$

Whenever $x \notin \mathcal{R}(\psi(H))$ such distance functions $d_\psi(R)$, which are positive, decreasing, convex and continuous for all $0 \leqslant R < \infty$, moreover tending to zero as $R \to \infty$, measure the degree of violation of x with respect to the benchmark source condition (4). Most of the mentioned properties of the function d_ψ are given in [10, lemma 2.5 and its proof]. For completeness we indicate the proof of the convexity. If for $0 < R, S < \infty$ the minimizers for $d_\psi(R)$ and $d_\psi(S)$ are called v^R, v^S, respectively, then $\|(v^R + v^S)/2\| \leqslant (R + S)/2$, and $d_\psi((R + S)/2) \leqslant \|x - \psi(H)((v^R + v^S)/2)\| \leqslant \frac{1}{2}\,(d_\psi(R) + d_\psi(S))$.

Remark 1. In the previous study [13] the authors used the distance function ϱ_ψ defined as

$$\varrho_\psi(t) := \inf\{\|tx - \psi(H)v\| : \ \|v\| \leqslant 1\}, \qquad t > 0,$$

instead of $d_\psi(R),\ R > 0$. The relation

$$d_\psi(R) = R\varrho_\psi(1/R), \qquad R > 0, \tag{7}$$

is evident from the definition. We know from [13, lemma 5.3] that the function ϱ_ψ as well as the function $t \mapsto \varrho_\psi(t)/t,\ t > 0$, are increasing index functions mapping $(0, \infty)$ onto itself. Some more insight gives [8, remark 1] indicating that the transformations $\varrho_\psi \mapsto d_\psi$ according to (7) and its inverse $d_\psi \mapsto \varrho_\psi$ are *involutions* that preserve convexity.

Here we confine to the usage of d_ψ as the distance function. In the following, for short we call the index function ψ in the d_ψ *benchmark*. If the decay of $d_\psi(R) \to 0$ as $R \to \infty$ is slow, then x strongly violates the benchmark source condition, whereas a fast decay corresponds to a weak violation. The distance function approach, also called *method of approximate source conditions* is a third way of expressing solution smoothness. To see this we recall the following result, where as in the following we write $d_\nu(R)$ for short if we mean $d_\psi(R)$ with the monomial benchmark $\psi(t) = t^\nu,\ \nu > 0$.

Fact 3 ([4, theorem 3.2], [13, section 5.2])**.** *Smoothness of the form $x \in \mathcal{R}(H^\kappa)$ with $0 < \kappa < \nu$ implies for the distance function with the monomial benchmark $\psi(t) = t^\nu,\ t > 0$, a decay rate $d_\nu(R) = \mathcal{O}\left(R^{-\frac{\kappa}{\nu-\kappa}}\right)$ as $R \to \infty$.*

This assertion will be improved as a result of our analysis in corollary 2.

We should mention here the dual formulation of the distance function

$$d_\psi(R) = \sup\left\{\langle x, v\rangle - R\|\psi(H)v\|,\ \|v\| \leqslant 1\right\}, \qquad R > 0, \tag{8}$$

as an alternative to (6), which is derived from the concept of the Fenchel duality, see [28, section 2.7]. This can be helpful for obtaining lower bounds of $d_\psi(R)$ if one finds an appropriate choice of $v \in X$, see for example [16, p 96].

Remark 2. Distance functions, or approximate source conditions, already mentioned in the monograph [1], are also used for the analysis of nonlinear ill-posed problems. If the nonlinear forward mapping is smoothing then the Fréchet derivative at the solution constitutes a linear operator, which is compact in typical cases. This can then be used to measure smoothness, and to establish corresponding distance functions. We mention [2, 9] for recent references concerning Tikhonov-type regularization, and [18] concerning iterative regularization. However, the error analysis of nonlinear equations is superimposed by the specific structure of nonlinearity in a neighborhood of the solution, which prevents an immediate use of the techniques as developed below. Nonetheless, a prospective analysis in this direction may result in new insights.

Outline. In this study we will establish relations between the different kinds of measuring solution smoothness, in particular between distance functions $d_\psi(R)$, the distribution function $F_x(\alpha)$, and a regularization error $f_x(\alpha)$ in the context of the linear regularization schema introduced in section 2. The basic question is, whether the distance function carries the same spectral information as the distribution function. General lower and upper bounds are given in section 3. The proof of the main general result will be given in the appendix. For the specific case of the power-type behavior we even show a one-to-one correspondence between the associated exponents in section 4, which answers the above basic question in the affirmative. We also highlight the theoretical results by providing examples which show the application and limitations of our findings in section 5. In particular we discuss relations to other results in this direction, which were previously obtained by several authors.

2. The linear regularization schema

Before formulating the main results in sections 3 and 4 we shall recall the concept of a regularization schema as given in [13], where only linear regularization operators R_α are under consideration. If the family of operators R_α was obtained from some family of generator functions g_α, see (2), then the regularization error constitutes as

$$x - R_\alpha A x = (I - g_\alpha(H)H)x,$$

and we associate to the function g_α the residual (bias) function $r_\alpha(t) := 1 - t g_\alpha(t),\ 0 < t \leqslant \|H\|$.

Definition 1. *A family of piecewise continuous functions $g_\alpha(t)$ is called a regularization if $\lim\limits_{\alpha\to 0} r_\alpha(t) = 0$ as $\alpha \to 0$ for all $0 < t \leqslant \|H\|$ and the following estimates hold for all $0 < t \leqslant \|H\|,\ 0 < \alpha \leqslant \bar{\alpha}$, and with constants γ_0, γ_1 and γ_*:*

(1) $\sqrt{t}\,|g_\alpha(t)| \leqslant \gamma_*/\sqrt{\alpha}$,
(2) $|r_\alpha(t)| \leqslant \gamma_1$,
(3) $t\,|g_\alpha(t)| \leqslant \gamma_0$.

Using the notation of index functions introduced in the initial section we say that an index function φ is a *qualification* of the regularization generated by g_α with constant γ if

$$|r_\alpha(t)|\,\varphi(t) \leqslant \gamma\varphi(\alpha), \qquad 0 < t \leqslant \|H\|, \qquad 0 < \alpha \leqslant \bar{\alpha}.$$

A standard account on the linear regularization schema is [5].

Regularization from a single function. In many cases the regularization family g_α can be obtained from a single function, say $g : (0,\infty) \to \mathbb{R}$, accompanied with the function $r(t) := 1 - t g(t),\ t > 0$, where we refer to [27, section 2.3] and to the German textbook [20]. It is easy to see that such function g gives rise to a regularization by exploiting the dilatation procedure:

$$g_\alpha(t) := \frac{1}{\alpha} g\left(\frac{t}{\alpha}\right), \qquad t > 0,\ \alpha > 0,$$

provided that g obeys

(1) $\sqrt{t}\,|g(t)| \leqslant \gamma_*,\quad t > 0$,
(2) $|r(t)| \leqslant \gamma_1,\quad t > 0$, and
(3) $t\,|g(t)| \leqslant \gamma_0,\quad t > 0$.

Note that by construction we have that $r_\alpha(t) = r(t/\alpha)$ for $t, \alpha > 0$, and that r_α is well defined for every $\alpha > 0$, i.e., $\bar{\alpha} := \infty$.

Profile functions. With respect to any given regularization g_α we are interested in the profile function (regularization error), which is seen to equal

$$f_x(\alpha) := \|r_\alpha(H)x\|, \qquad 0 < \alpha \leqslant \bar{\alpha}. \tag{9}$$

We shall assume that $\|r_\alpha(H)x\|$ is a non-decreasing function in α for simplicity. This is the case if the underlying regularization g_α is such that the function $\alpha \mapsto |r_\alpha(t)|,\ 0 < \alpha \leqslant \bar{\alpha}$, is non-decreasing, which is always satisfied for regularization from a single function g with non-increasing r. If the regularization g_α consists of continuous functions, then f_x is continuous and hence with $\lim_{\alpha\to 0} f_x(\alpha) = 0$ an index function.

Example 1 (Tikhonov regularization)**.** This regularization method with the continuous functions $g_\alpha(t) = 1/(t+\alpha)$ and $r_\alpha(t) = \alpha/(t+\alpha)$ is obtained from $g(t) := 1/(t+1)$, with $r(t) = 1/(t+1)$. We observe that the function r is decreasing and moreover that $\frac{1}{2} = r(1) \leqslant r(t) \leqslant 1,\ 0 < t \leqslant 1$. Any concave index function, in particular any linear, is a qualification of the method.

Example 2 (spectral cutoff)**.** Another important regularization method is spectral cutoff, where $g_\alpha(t) = 1/t,\ t > \alpha$, and $g_\alpha(t) = 0$, otherwise. This function with a jump at the point $t = \alpha$ corresponds to $g(t) = 1/t,\ t > 1$, and $g(t) = 0$, otherwise. The residual function is $r(t) = \chi_{(0,1]}(t)$, the characteristic function of the interval $(0, 1]$, again a non-increasing function in t. Spectral cutoff has arbitrary index functions as qualification. We note the important observation that for spectral cutoff we have that

$$\|r_\alpha(H)x\| = F_x(\alpha), \qquad 0 < \alpha \leqslant \|H\|. \tag{10}$$

Remark 3. It is well known that regularization can also be achieved by discretization, or a combination of linear regularization with discretization, we mention [23] or the more recent [19]. Here the regularization error measures the behavior of the discretization on exact data. This is also an important quantity, and the smoothness concepts as developed here have impact on the decay of the regularization error through the degree of approximation and the modulus of injectivity (inverse property). Details are beyond the scope of this study, we refer to [14] for some initial study in this direction, even for nonlinear equations.

One important observation from [24] relates the distribution function F_x to the profile function f_x as follows.

Proposition 1. *For every regularization* g_α *there is a constant* $0 < c \leqslant 1$ *such that*

$$F_x(c\alpha) \leqslant 2 f_x(\alpha), \qquad 0 < \alpha \leqslant \|H\|.$$

Proof. We first note that by item (i) of definition 1 with constant γ_* we find a $0 < c \leqslant 1$ for which $|r_\alpha(t)| \geqslant 1/2$ whenever $0 < t \leqslant c\alpha$. Precisely, setting $c := \min\left\{\frac{1}{4\gamma_*^2}, 1\right\}$ we have

$$|r_\alpha(t)| \geqslant 1 - t\,|g_\alpha(t)| \geqslant 1 - \gamma_*\sqrt{t/\alpha} \geqslant 1 - \gamma_*\sqrt{c} \geqslant 1/2.$$

Then we bound

$$\begin{aligned} f_x^2(\alpha) &= \int_0^{\|H\|} |r_\alpha(t)|^2 \;\mathrm{d}F_x^2(t) \\ &\geqslant \int_0^{c\alpha} |r_\alpha(t)|^2 \;\mathrm{d}F_x^2(t) \geqslant \frac{1}{4}\int_0^{c\alpha} \;\mathrm{d}F_x^2(t) = \frac{1}{4}F_x^2(c\alpha), \end{aligned}$$

from which the assertion follows. □

3. General results

As already mentioned, both functions, the regularization error f_x and the distance function d_ψ reflect smoothness of x in the sense of certain spectral properties of the involved element with respect to H.

We shall establish for $x \notin \mathcal{R}(\psi(H))$ a one-to-one correspondence between distance functions $d_\psi(R)$ and profile functions $f_x(\alpha)$, for an appropriate relation between R and α. To this end let

$$\Phi_\psi(R) := \frac{d_\psi(R)}{R} = \varrho_\psi\left(\frac{1}{R}\right), \qquad R > 0. \tag{11}$$

Since $d_\psi(R)$ is decreasing and continuous for $0 \leqslant R < \infty$, the function $\Phi_\psi(R)$ is even a strictly decreasing continuous function for all positive R mapping $(0, \infty)$ onto itself. By the above reasoning the equation

$$\Phi_\psi(R) = \psi(\alpha) \tag{12}$$

has a unique solution $R = R(\alpha)$ for each $\alpha \in (0, \bar{\alpha}]$.

The following upper bound was derived by using the function ϱ_ψ in [13, theorem 5.5]. For the convenience of the reader we recall the proof within the present context, here.

Proposition 2. *If the regularization g_α has qualification ψ with constant γ and if $x \notin \mathcal{R}(\psi(H))$, then*

$$\|r_\alpha(H)x\| \leqslant (\gamma + \gamma_1) d_\psi\left(\Phi_\psi^{-1}(\psi(\alpha))\right), \qquad 0 < \alpha \leqslant \bar{\alpha}.$$

Proof. Let $R = R(\alpha)$ be given by solving equation (12), and let v be the minimizer of the distance function $d_\psi(R)$. Then

$$\begin{aligned}\|r_\alpha(H)x\| &\leqslant \|r_\alpha(H)(x - \psi(H)v)\| + \|r_\alpha(H)\psi(H)v\| \\ &\leqslant \gamma_1 d_\psi(R) + \gamma\psi(\alpha)R \\ &\leqslant (\gamma_1 + \gamma)\max\{d_\psi(R), \psi(\alpha)R\},\end{aligned}$$

which allows us to complete the proof. □

Our goal is to establish a converse to the bound from proposition 2, and we shall distinguish a *low-benchmark* and a *high-benchmark* case, respectively. The main general result is the following theorem 1. We establish essential cross connections between distance and profile functions. This will be helpful for discussing the power-type case in the subsequent section. Because the proof of this theorem is rather technical, we postpone it to the appendix section.

Theorem 1. *Let ψ be an index function and $x \notin \mathcal{R}(\psi(H))$. Moreover let g_α be a regularization.*

***Low-benchmark**: If the function $t \mapsto \psi^2(t)/t$ is non-increasing, or*
***High-benchmark**: If the function $t \mapsto \psi^2(t)/t$ is non-decreasing, and the regularization is obtained from a single function g with non-increasing r, $|r(1)| > 0$,*

then

$$d_\psi\left(2\max\{\gamma_0, \gamma_*\}\Phi_\psi^{-1}(\psi(\alpha))\right) \leqslant C\|r_\alpha(H)x\|, \qquad 0 < \alpha \leqslant \bar{\alpha}, \tag{13}$$

with constant $C = 1$ in the low-benchmark case and $C = \gamma_1/|r(1)|$ in the high-benchmark case.

If the regularization has qualification ψ, then

$$\|r_\alpha(H)x\| \leqslant 2\max\{\gamma, \gamma_1\} d_\psi\left(\Phi_\psi^{-1}(\psi(\alpha))\right), \qquad 0 < \alpha \leqslant \bar{\alpha}.$$

Remark 4. We mention that $\gamma_0 \geqslant 1$ and hence $2\max\{\gamma_0, \gamma_*\} \geqslant 2$ for any regularization g_α, and this follows from $\lim\limits_{\alpha\to 0} t\, g_\alpha(t) = 1$ for all $0 < t \leqslant \|H\|$ as required in definition 1, which implies that $\gamma_0 \geqslant \sup\limits_{0<t\leqslant\|H\|,\ 0<\alpha\leqslant\bar{\alpha}} t\,|g_\alpha(t)| \geqslant 1$.

In case that the distance function d_ψ does not decay too quickly we can take out the leading constant $2\max\{\gamma_0, \gamma_*\}$ on the left in (13). Precisely, a non-decreasing function, say $h : (0, \infty) \to (0, \infty)$, is said to obey a Δ_2-condition if there is a constant $C_2 \geqslant 1$ such that $h(2t) \leqslant C_2 h(t)$, $t > 0$. This restricts the growth rate of the function h to be sub-exponential.

Within the present context this specifies to the following. If the function $1/d_\psi$ obeys a Δ_2-condition, in particular, if d_ψ decays sub-exponentially, then there is a constant $0 < c \leqslant 1$ such that

$$cd_\psi\big(\Phi_\psi^{-1}(\psi(\alpha))\big) \leqslant d_\psi\big(2\max\{\gamma_0, \gamma_*\}\Phi_\psi^{-1}(\psi(\alpha))\big).$$

In this case the upper and lower bounds for the profile function in theorem 1 coincide up to constants.

We highlight one specific instance of theorem 1, when we choose spectral cutoff as regularization. In this case the profile function $f_x(\alpha)$ has a clear interpretation, see representation (10) in example 2.

Corollary 1. *For an arbitrary index function ψ we have that*

$$d_\psi\big(2\Phi_\psi^{-1}(\psi(\alpha))\big) \leqslant F_x(\alpha) \leqslant 2d_\psi\big(\Phi_\psi^{-1}(\psi(\alpha))\big), \qquad 0 < \alpha < \infty.$$

Proof. The proof of the lower bound is the same as for the high-benchmark case of theorem 1 applied to the spectral cut-off schema, but instead of lemma 8 here one has to use remark 7, see the appendix below. For the upper bound we apply theorem 1 again to spectral cutoff and take into account that $\gamma_* = \gamma_0 = 1$. □

We close this section with general remarks on techniques for bounding distance functions. The 'standard' way for establishing upper bounds is to use the family $w_\alpha := \chi_{(\alpha,\|H\|]}(H)\psi(H)^{-1}x \in X,\ 0 < \alpha \leqslant \bar{\alpha}$, since here $\|x - \psi(H)w_\alpha\| = F_x(\alpha)$. This family adapts the choice used for proving lemma 1. Upper bounds are then obtained from this, since $d_\psi(\|w_\alpha\|) \leqslant F_x(\alpha)$, such that upper bounds for $\|w_\alpha\|$, typically of the order $F_x(\alpha)/\psi(\alpha)$, yield upper bounds for d_ψ by monotonicity.

It is harder to establish lower bounds. Two ways are worth mentioning. First, corollary 1 yields, with $R := \Phi_\psi^{-1}(\psi(\alpha))$ that

$$F_x(\psi^{-1}(\Phi_\psi(R))) \leqslant 2d_\psi(R), \qquad R \geqslant R_0.$$

In the 'moderate' cases, if the smoothness given by the distribution F_x is far enough from the benchmark ψ, this yields lower bounds for d_ψ, see e.g. the reasoning in example 3(c), below. However, in extremal cases, if the actual smoothness is close to the benchmark, then a case-dependent analysis may provide sharp lower bounds. Here we point at the corresponding discussion in example 3(a).

4. Power-type behavior

Here we discuss consequences of corollary 1 for power-type functions of growth and decay rates expressing the solution smoothness. In general such functions correspond to moderate smoothness situations; other behavior is possible, see example 4. In the following result we consider power-type benchmark smoothness $\psi_\nu(t) := t^\nu$ for some fixed $\nu > 0$ with the corresponding distance functions $d_\nu(R) := d_{\psi_\nu}(R)$ and associated quotient functions $\Phi_\nu(R) := \frac{d_\nu(R)}{R}$.

Since the results will be asymptotic in nature, we recall the following notion and notation.

Definition 2. *Suppose that $f, g : (0, a) \to (0, \infty)$ are real functions. Then we denote $f = \mathcal{O}(g)$ as $t \to 0$, if there are constants $C < \infty$ and $0 < \bar{t} \leqslant a$ such that $f(t) \leqslant Cg(t),\ 0 < t < \bar{t}$. We denote $f \asymp g$ if $f = \mathcal{O}(g)$ and $g = \mathcal{O}(f)$. Finally, we denote $f = o(g)$ if $f(t)/g(t) \to 0$ as $t \to 0$.*

The above behavior is concerned with functions defined in a right neighborhood of zero, but similar notion and notation applies for positive functions $f(R),\ g(R),\ R \in [M, \infty),\ M > 0$, when the limit case $R \to \infty$ is under consideration.

The following result extends and reproves [24, theorem 2.2] by using distance functions. Its proof, as given here, will use the results obtained so far.

Theorem 2. *Let $0 < \kappa < \nu$. For $x \in X$ satisfying the condition $x \notin \mathcal{R}(H^\nu)$ the following assertions are equivalent.*

(1) The distribution function for x behaves like $F_x(\alpha) \asymp \alpha^\kappa$ as $\alpha \to 0$.
(2) The distance function for x behaves like $d_\nu(R) \asymp R^{-\frac{\kappa}{\nu-\kappa}}$ as $R \to \infty$.
(3) For an arbitrary regularization g_α, which has $\varphi(t) = t^\nu,\ t > 0$, as qualification, the profile function $f_x(\alpha) := \|r_\alpha(H)x\|$ for x behaves like $f_x(\alpha) \asymp \alpha^\kappa$ as $\alpha \to 0$.

The assertions remain valid when replacing $\asymp$ by either big-o '$\mathcal{O}$' or little-o 'o'.

We shall see in the examples 3 and 5, given in section 5, that the power-type behavior actually occurs for compact and non-compact operators H.

To prove theorem 2 we start with the following preliminary result.

Lemma 2. *Let $0 < \kappa < \nu$. We have that $d_\nu(R) = \mathcal{O}(R^{-\kappa/(\nu-\kappa)})$ as $R \to \infty$ implies that $d_\nu(\Phi_\nu^{-1}(\alpha)) = \mathcal{O}(\alpha^\kappa)$ as $\alpha \to 0$. The assertion remains true if we replace big-O '$\mathcal{O}$' by little-o 'o'.*

Proof. The assumption on d_ν implies that $\Phi_\nu(R) = \mathcal{O}(R^{-1/(\nu-\kappa)})$; thus, there are $0 < C < \infty$ and $0 < R_0 < \infty$ for which $\Phi_\nu(R) \leqslant CR^{-1/(\nu-\kappa)}$ whenever $R \geqslant R_0$. This yields that $\Phi_\nu^{-1}(CR^{-1/(\nu-\kappa)}) \geqslant R$, by monotonicity. We assign $\alpha := CR^{-1/(\nu-\kappa)}$ and thus have that $\Phi_\nu^{-1}(\alpha) \geqslant (\alpha/C)^{\kappa-\nu}$. Therefore, using the monotonicity of d_ν we see that

$$d_\nu\left(\Phi_\nu^{-1}(\alpha)\right) \leqslant d_\nu((\alpha/C)^{\kappa-\nu}) \leqslant C((\alpha/C)^{\kappa-\nu})^{-\kappa/(\nu-\kappa)} = \mathcal{O}(\alpha^\kappa).$$

The assertion for little-o is along the same lines, but more tedious. □

Proof of theorem 2. We first proof the equivalence of items (I) and (II) in either of the asymptotic regimes. By lemma 2 and corollary 1 the order of magnitude of d_ν yields the corresponding order for F_x. For the converse we use the approximation $w_\alpha := \chi_{(\alpha,\|H\|]}(H)H^{-\nu}x$. By substituting the operator H in lemma 1 with H^ν we see that $F_x(\alpha) = \mathcal{O}(\alpha^\kappa)$ yields an inequality of the form $\alpha^\nu\|w_\alpha\| = \mathcal{O}(\alpha^\kappa)$. Therefore, there is $0 < C < \infty$ such that

$$d_\nu(C\alpha^{\kappa-\nu}) \leqslant d_\nu(\|w_\alpha\|) \leqslant \|x - H^\nu w_\alpha\| = F_x(\alpha).$$

By letting $\alpha := (R/C)^{-1/(\nu-\kappa)}$ we obtain that

$$d_\nu(R) \leqslant F_x((R/C)^{-1/(\nu-\kappa)}) = \mathcal{O}(R^{-1/(\nu-\kappa)}) \text{ as } R \to \infty.$$

Again, for little-o the reasoning is similar.

Finally, suppose that $F_x(\alpha) \asymp \alpha^\kappa$. Plainly, by the first part of the proof this implies that $d_\nu(R) = \mathcal{O}(R^{-\kappa/(\nu-\kappa)})$. If it were true that $d_\nu(R) = o(R^{-\kappa/(\nu-\kappa)})$, then this would imply, again by the beginning of the proof, that $F_x(\alpha) = o(\alpha^\kappa)$, contradicting the assumption. Similar applies by assuming that $d_\nu(R) \asymp R^{-\kappa/(\nu-\kappa)}$, and this shows the equivalence of items (I) and (II) in either of the asymptotic regimes.

Next, the asymptotics $d_\nu(R) = \mathcal{O}(R^{-\frac{\kappa}{\nu-\kappa}})$ implies that $f_x(\alpha) = \mathcal{O}(\alpha^\kappa)$, by lemma 2 and proposition 2. The same applies for little-o 'o'.

Finally, if $F_x(\alpha) \asymp \alpha^\kappa$ then, by the first part of the proof, and by proposition 2, this implies that $f_x(\alpha) = \mathcal{O}(\alpha^\kappa)$. From proposition 1 we deduce that $\alpha^\kappa = \mathcal{O}(f_x(\alpha))$ in this case.

Since item (I) is a special case of item (III), the proof is complete. □

This shows that in power-type situations, the decay rates of the distance function as well as of the profile function both relate to the 'pointwise' behavior of the spectrum of x with respect to H, given by F_x. In contrast smoothness expressed by source conditions (4), i.e. requiring that x is in the range of some index function of the operator H, considers the spectrum of H in an 'integral' form only. We dwell on this, and highlight some relations to the study [4].

Lemma 3. *Let ψ be an arbitrary index function. If $x \in \mathcal{R}(\psi(H))$ then $F_x(\alpha) = o(\psi(\alpha))$ as $\alpha \to 0$.*

Proof. First, by the injectivity of the operator H we have $F_x^2(\alpha) \to 0$ as $\alpha \to 0$. Next, by assumption the function $1/\psi^2 \in L_1((0,\infty), \mathrm{d}F_x^2)$, and hence the measure $dG_x(t) := 1/\psi^2(t)\,\mathrm{d}F_x^2(t)$ is absolutely continuous with respect to $\mathrm{d}F_x^2$. Thus, for each $\varepsilon > 0$ there is $\delta > 0$ such that $\int_{\mathcal{A}} \mathrm{d}F_x^2(t) \leqslant \delta$ implies that $\int_{\mathcal{A}} dG_x(t) \leqslant \varepsilon$ for all Borel sets $\mathcal{A} \subset (0,\infty)$. In particular, if α_δ is small enough such that $F_x^2(\alpha_\delta) \leqslant \delta$, then

$$\varepsilon \geqslant \int_0^{\alpha_\delta} \frac{1}{\psi^2(t)}\,\mathrm{d}F_x^2(t) \geqslant \frac{1}{\psi^2(\alpha_\delta)} F_x^2(\alpha_\delta),$$

from which the proof can be completed. □

In the light of theorem 2 this yields the following strengthening and generalization of [4, theorems 3.1 and 3.2].

Corollary 2. *Let $0 < \eta < \nu$, and suppose that $x \in \mathcal{R}(H^\eta)$ but $x \notin \mathcal{R}(H^\nu)$. Then*

$$d_\nu(R) = o(R^{-\eta/(\nu-\eta)}) \quad \text{as } R \to \infty, \tag{14}$$

and

$$f_x(\alpha) = o(\alpha^\eta) \quad \text{as } \alpha \to 0, \tag{15}$$

for an arbitrary regularization g_α, which has $\varphi(t) = t^\nu,\ t > 0$, as qualification.

Proof. By lemma 3 we deduce from the assumption that $F_x(\alpha) = o(\alpha^\eta)$. Now, theorem 2, in the little-o 'o' cases, yields both (14) and (15). □

This explains that optimal decay rates for the regularization error based on general smoothness in terms of index functions ψ in (4) cannot be obtained, as this was already mentioned after lemma 1. This can also be seen from [22]. In addition, we refer to the discussion on lower bounds in [13], and also the study [21, section 4].

We finally mention that sharp bounds for the profile function f_x yield sharp bounds for the overall error, and this was emphasized in [25, theorem 2.6]. In particular, for Tikhonov regularization as in remark 1, given noisy measurements $y^\delta = Ax + \delta\xi,\ \|\xi\| \leqslant 1$, we assign the family of approximate solutions $x_\alpha^\delta := R_\alpha y^\delta = (A^*A + \alpha I)^{-1}A^*y^\delta,\ \alpha > 0$. Then, in any of the cases (I)–(III) in theorem 2 the order of reconstruction of x from noisy measurements obeys

$$\sup_\xi \left\{ \inf_{\alpha>0} \|x_\alpha^\delta - x\|,\ \|Ax - y^\delta\| \leqslant \delta \right\} \asymp \delta^{\frac{\kappa}{\kappa+1/2}}, \qquad \delta \to 0.$$

Thus, there are always noise instances ξ such that one cannot beat the rate $\delta^{\kappa/(\kappa+1/2)}$ by any means of choice of the regularization parameter α. It is also pointed out in corollary 2.8 ibid. that any reconstruction rate $\delta^{\mu/(\mu+1/2)}$ necessarily yields that $x \in \mathcal{R}((A^*A)^\varrho)$ for all $0 < \varrho < \mu$. We will not dwell into this, but refer to [25] for further details.

5. Examples

We turn to discussing examples in connection with the theoretical results. The following lemma will be useful.

Lemma 4. *Let ψ be any index function and $x \in X$ any element. For $\lambda > 0$ let $u_\lambda := (\lambda + \psi^2(H))^{-1}\psi(H)x$. Then*

$$d_\psi(\|u_\lambda\|) = \lambda\|(\lambda + \psi^2(H))^{-1}x\|, \qquad \lambda > 0.$$

Proof. Set $R := \|u_\lambda\|$. The element u_λ obeys the equation

$$\psi(H)(\psi(H)u_\lambda - x) = -\lambda u_\lambda,$$

and it is thus a minimizer of $\|\psi(H)w - x\|^2$ with a constraint of the form $\|w\| \leqslant R$. Therefore,

$$d_\psi(\|u_\lambda\|) = \|\psi(H)u_\lambda - x\| = \lambda\|(\lambda + \psi^2(H))^{-1}x\|,$$

which proves the assertion. □

The first example is based on a non-compact multiplication operator, in particular it points at the limitations of the bounds from theorem 1 when the distance functions do not obey some Δ_2-condition.

Example 3. The authors in [10, 12] have discussed injective multiplication operators in $L_2(0, 1)$ with non-closed range, and [12, example 4.6], or [10, section 3, example 2], will guide us, here. Precisely, we focus on the non-compact operator

$$[Hx](s) := m(s)x(s), \qquad 0 < s < 1, \tag{16}$$

with the multiplier function $m \in L^\infty(0, 1)$ possessing an essential zero. Below, we shall restrict to cases where m is strictly increasing and continuous for $0 < s < 1$ with limit conditions $\lim_{t\to 0} m(t) = 0$ and $\lim_{t\to 1} m(t) = 1$. It is evident for such multiplication operators that

$$F_x^2(t) = \int_{\{s\in(0,1):\, 0<m(s)\leqslant t\}} x^2(s)\,\mathrm{d}s = \int_0^{m^{-1}(t)} x^2(s)\,\mathrm{d}s, \qquad 0 < t \leqslant 1. \tag{17}$$

This distribution function is continuous with $F_x(1) = \|x\|$ and can be extended continuously as $F_x(t) = \|x\|$ for all $1 < t < \infty$. So we can easily derive the following representation for an arbitrary function $h \in L_2((0,\infty), \mathrm{d}F_x^2)$, namely that

$$\|h(H)x\|^2 = \int_0^\infty h^2(t)\,\mathrm{d}F_x^2(t) = \int_0^1 h^2(m(s))x^2(s)\,\mathrm{d}s. \tag{18}$$

We now restrict to the case that $x \equiv 1$ is a constant function; hence, $F_x^2(t) = m^{-1}(t),\ 0 < t \leqslant 1$. We note that $1 \in \mathcal{R}(H)$ if and only if $1/m \in L_2(0, 1)$, and hence we shall confine ourselves to distance functions $d_1(R),\ R > 0$, with respect to the benchmark $\psi(t) = t,\ t > 0$.

We will derive results for the quantities under consideration for different multipliers m for which $1 \notin \mathcal{R}(H)$. In our considerations we shall need the family $w_\alpha := \chi_{(\alpha,1]}(H)H^{-1}1$, for which $1 - Hw_\alpha = \chi_{(0,\alpha]}(H)1$, thus $\|1 - Hw_\alpha\| = F_x(\alpha)$. Also, representation (18) yields that

$$\|w_\alpha\|^2 = \int_{m^{-1}(\alpha)}^1 \frac{1}{m^2(s)}\,\mathrm{d}s. \tag{19}$$

The following cases for multipliers are of interest.

- $m(s) = \sqrt{s}$: this is the limiting case for $1 \notin \mathcal{R}(H)$. Here we have $F_x(\alpha) = \alpha,\ 0 < \alpha \leqslant 1$ and hence $F_x(\alpha) \asymp \alpha$, i.e. item (I) in theorem 2 is satisfied with $\kappa = 1$, but we cannot apply this theorem with $\nu = 1$ and benchmark $\psi(t) = t,\ t > 0$, since the required strict inequality $\kappa < \nu$ fails. This is reflected in the following behavior of the distance function, and we shall use lemma 4 with function $\psi(t) = t,\ t > 0$, which gives $d_1(g(\lambda)) = h(\lambda),\ \lambda > 0$ with

$$g(\lambda) := \|u_\lambda\| = \sqrt{\int_0^1 \frac{t}{(\lambda+t)^2}\,\mathrm{d}t} = \sqrt{\ln\frac{\lambda+1}{\lambda} - \frac{1}{\lambda+1}},$$

$$\textit{and} \quad h(\lambda) := \lambda\|(\lambda+H^2)^{-1}x\| = \sqrt{\int_0^1 \frac{\lambda^2}{(\lambda+t)^2}\,\mathrm{d}t} = \sqrt{\frac{\lambda}{\lambda+1}}.$$

The re-parametrization $u := \frac{\lambda}{\lambda+1} \in (0,1)$ yields $d_1(\tilde{g}(u)) = \sqrt{u}$, for the decreasing function $\tilde{g}(u) := \sqrt{u - \ln u - 1},\ u \in (0,1)$. First, for $u \in (0,1)$ we have $-\ln u > -\ln u + u - 1$, and we obtain that

$$d_1(\sqrt{-\ln u}) \leqslant d_1(\sqrt{-\ln u + u - 1}) = \sqrt{u}, \qquad 0 < u < 1.$$

For $u := e^{-R^2}$ this results in the upper bound $d_1(R) \leqslant e^{-R^2/2}$ for $R > 0$.
For the lower bound, given $R > 0$ we find u_R with $R := \tilde{g}(u_R)$, and hence $d_1(R) = d_1(\tilde{g}(u_R)) = \sqrt{u_R}$. From $\tilde{g}(e^{-(R^2+1)}) > R$ we conclude that $u_R > e^{-(R^2+1)}$, hence $d_1(R) > e^{-(R^2+1)/2} = e^{-1/2}e^{-R^2/2},\ R > 0$. Thus, $d_1(R) \asymp e^{-R^2/2}$ as $R \to \infty$, expressing a very high decay rate of the distance function.
This yields, for some unspecified constants $0 < \underline{c} \leqslant \overline{c} < \infty$, and for $c := (2\max\{\gamma_0, \gamma_*\})^2 \geqslant 4$, see remark 4, that

$$\underline{c}\alpha^c \leqslant \|r_\alpha(H)1\| \leqslant \overline{c}\sqrt{\alpha^2 \log\left(\frac{1}{\alpha^2}\right)}, \qquad 0 < \alpha \leqslant \bar{\alpha}. \tag{20}$$

Indeed, for the upper estimate we observe $\Phi_1(R) \leqslant e^{-R^2/2}$ for $R \geqslant 1$. Thus, $\Phi_1^{-1}(\alpha) \leqslant \sqrt{\ln\frac{1}{\alpha^2}}$ for α small enough, and the upper bound in theorem 1 gives

$$\|r_\alpha(H)1\| \leqslant \overline{c}d_1\left(\Phi_1^{-1}(\alpha)\right) = \overline{c}\alpha\Phi_1^{-1}(\alpha) \leqslant \overline{c}\sqrt{\alpha^2\ln\frac{1}{\alpha^2}}.$$

By similar arguments we obtain the lower bound:

$$\|r_\alpha(H)1\| \geqslant \tilde{c}d_1\left(\sqrt{c}\Phi_1^{-1}(\alpha)\right) \geqslant \tilde{c}d_1\left(\sqrt{c\ln\frac{1}{\alpha^2}}\right) \geqslant \tilde{c}\exp\left(-\frac{1}{2}\left(c\ln\frac{1}{\alpha^2}+1\right)\right)$$
$$= \tilde{c}\exp(-\tfrac{1}{2})\alpha^c.$$

We observe from (20) that theorem 1 does not give precise bounds for the profile function. For Tikhonov regularization we can derive the behavior of the profile function, and it behaves like $f_x(\alpha) \asymp \sqrt{\alpha^2\log(\frac{1}{\alpha^2})}$ as $\alpha \to 0$. Indeed, in view of the upper bound in (20), we only need a lower bound, and this can be derived as

$$f_x^2(\alpha) = \|r_\alpha(H)1\|^2 = \int_0^1 \frac{\alpha^2}{(\sqrt{s}+\alpha)^2}\,\mathrm{d}s \geqslant c\alpha^2\log(1/\alpha),$$

for α small enough and some constant $c > 0$, such that the upper bound in (20) has the right order of magnitude. Note also the different rates of increase for the distribution function ($F_x(\alpha) \asymp \alpha$) and the profile function.

Table 1. This table summarizes the asymptotic behavior of the quantities of interest in cases (a)–(c). Note that the upper and lower bound for $f_x(\alpha)$ in the case $m(s) = \sqrt{s}$ differ.

$m(s)$	$F_x(\alpha)$	$f_x(\alpha)$	$d_1(R)$
$\sqrt{s}$	α	$\alpha^c, \sqrt{\alpha^2 \log(\frac{1}{\alpha^2})}$	$e^{-R^2/2}$
s	$\sqrt{\alpha}$	$\sqrt{\alpha}$	$1/R$
$e^{1-1/\sqrt{s}}$	$1/\log(1/\alpha)$	$1/\log(1/\alpha)$	$1/\log(R)$

- $m(s) = s$: this is an intermediate case, and we shall see that the main results are sharp, here. In this case $F_x(\alpha) = \sqrt{\alpha}$, $0 < \alpha \leqslant 1$, and by applying theorem 2 with $\nu = 1$ and $\kappa = 1/2$ we immediately find both, the corresponding rates for the distance function d_1 as
$$\underline{c}\,\frac{1}{R} \leqslant d_1(R) \leqslant \overline{c}\,\frac{1}{R}, \qquad 0 < R_0 \leqslant R < \infty,$$
as well as for the profile functions f_x as
$$\tilde{c}\,\sqrt{\alpha} \leqslant \|r_\alpha(H)1\| \leqslant \hat{c}\,\sqrt{\alpha}, \qquad 0 < \alpha \leqslant \bar{\alpha}.$$
- $m(s) = e^{1-1/\sqrt{s}}$: this is an example of the other limiting case, where theorem 2 also cannot be applied. Such multiplier function m results in a logarithmic increase for the distribution function, since $m^{-1}(\alpha) = 1/(1+\log(1/\alpha))^2$, and hence $F_x(\alpha) \asymp 1/\log(1/\alpha)$ as $\alpha \to 0$. To bound the distance function from above we use (19) to see that $\|w_\alpha\| \leqslant 1/\alpha$. Thus, $d_1(R) \leqslant 1/\log(R)$, $R > 0$. The right-hand side in corollary 1 yields a lower bound for the distance function. Indeed, this gives $\frac{c}{\log(R/d_1(R))} \leqslant d_1(R)$, and thus, since $\log(\xi) \leqslant c\xi/2$, $\xi \geqslant 2/c$,
$$\frac{c}{2}\frac{1}{d_1(R)} \leqslant c\frac{1}{d_1(R)} - \log\left(\frac{1}{d_1(R)}\right) \leqslant \log(R), \qquad R \geqslant R_0,$$
such that $d_1(R) \asymp 1/\log(R)$, which expresses a very low decay rate of the distance function as $R \to \infty$. On the other hand, the function $1/d_1$ obeys a Δ_2-condition in this case, and the bounds from theorem 1, and corollary 1, provide the right order of magnitude, such that also in this case we have the exact asymptotics for all quantities of interest.

  We summarize the derived asymptotic results in table 1.

Example 4. We use once more the multiplication operator (16), now with $m(s) = \sqrt{s}$, but we consider the function
$$x(s) := \sqrt{s^{-1/2}\log(1/s)}, \qquad 0 < s < 1.$$
By using (17), and partial integration, we see that
$$F_x^2(t) = \int_0^{t^2} s^{-1/2}\log(1/s)\,\mathrm{d}s = 2t\log(1/t^2) + 2\int_0^{t^2} s^{-1/2}\mathrm{d}s \asymp t\log(1/t),$$
as $t \to 0$. If we now fix the benchmark $\psi(t) = t$, then we see that $d_1(R) \asymp R^{-1}\log(R)$. This is a non-polynomial, but still moderate, behavior. Let us consider the quantity
$$\eta_{\sup} := \sup\{\eta > 0 : \; d_1(R) = \mathcal{O}(R^{-\eta/(1-\eta)})\} \in (0, 1], \tag{21}$$
and we see that for x from above it holds $\eta_{\sup} = 1/2$, but the supremum is *not attained* here. On the other hand, it still holds that $x \in \mathcal{R}(H^\mu)$ for all $\mu < 1/2$. Revisiting [4, corollary

3.3] we must note that the one-to-one correspondence of exponents in distance function and Hölder-type source conditions mentioned there in corollary 3.3, ibid., has to be rendered more precisely as $\eta_{sup} = \sup\{\mu > 0 : \ x \in \mathcal{R}(H^\mu)\}$.

Example 5. The authors in [16] studied the compact integration operator A in $L_2[0, 1]$, given as

$$[Ax](t) := \int_0^t x(\tau)\, \mathrm{d}\tau, \qquad 0 \leqslant t \leqslant 1.$$

These authors computed the asymptotics of the distance function for the constant function $x \equiv 1$ with benchmark A^*. Since $\mathcal{R}(A^*) = \mathcal{R}(H^{1/2})$ this corresponds to $\psi_{1/2}(t) = \sqrt{t}$, and it was proved there that $d_{1/2}(R) \asymp 1/R$ as $R \to \infty$. Within the present context this can be seen by evaluating the asymptotics of F_x, since we know the singular system $\{\sigma_i, u_i, v_i\}_{i=1}^\infty$ of the integration operator A with singular values $\sigma_i = \frac{1}{\pi(i-1/2)} \asymp i^{-1}$ and singular functions $u_i(t) = \sqrt{2}\cos((i-1/2)\pi t)$, $0 \leqslant t \leqslant 1,\ i = 1, 2, \ldots$. Then $\langle 1, u_i\rangle \asymp i^{-1}$ and consequently

$$\sum_{i=k}^\infty \langle 1, u_i\rangle^2 \asymp \sum_{i=k}^\infty \frac{1}{i^2} \asymp \frac{1}{k} \asymp \sigma_k \qquad \text{as} \qquad k \to \infty.$$

We know from [25, remark 2.2] that for all $\mu > 0$ the asymptotics $\sum_{i=k}^\infty \langle x, u_i\rangle^2 \asymp \sigma_k^{4\mu}$ is equivalent to $F_x(\alpha) \asymp \alpha^\mu$. Here we apply this for $\mu = 1/4$ and $x \equiv 1$. Then we have $F_x(\alpha) \asymp \alpha^{1/4}$, and we can use theorem 2 with $\kappa = 1/4$ and $\nu = 1/2$ to see that $f_x(\alpha) \asymp \alpha^{1/4}$, and also that $d_{1/2}(R) \asymp 1/R$, which concisely reproves the result from [16].

Acknowledgments

JF and BH were supported by the DFG under grant HO 1454/7-2.

Appendix. Proof of theorem 1

We first provide the necessary ingredients to prove theorem 1. In our followup analysis we will approximate the unknown minimizing element v in the definition of $d_\psi(R)$ within some suitable family v_α and w_α, respectively, and we introduce these, here. We (formally) assign

$$v_\alpha := g_\alpha(\psi^2(H))\psi(H)x, \tag{A.1}$$

and

$$w_\alpha := g_\alpha(H)H\psi(H)^{-1}x, \qquad 0 < \alpha \leqslant \bar{\alpha}, \tag{A.2}$$

where in case of w_α we have to assume that the operator $g_\alpha(H)H\psi(H)^{-1}$ is a bounded one. For spectral cutoff the element w_α is finite without any constraint on the function ψ, since there $w_\alpha = \int_\alpha^\infty \psi(t)^{-1}\, \mathrm{d}E_t x$ for $\alpha > 0$. Otherwise, for general regularization we provide the following sufficient conditions.

Lemma 5. *If the function $t \mapsto \psi^2(t)/t$ is non-increasing, then*

$$\|g_\alpha(H)H\psi(H)^{-1}\| \leqslant \frac{\max\{\gamma_0, \gamma_*\}}{\psi(\alpha)}, \qquad 0 < \alpha \leqslant \bar{\alpha}.$$

Thus, the element w_α is well defined in this case.

Proof. It is enough to show that $t\,|g_\alpha(t)|\,/\psi(t) \leqslant \max\{\gamma_0, \gamma_*\}/\psi(\alpha)$. If $t > \alpha$ then the monotonicity of ψ allows us to conclude that

$$t\,|g_\alpha(t)|\,/\psi(t) \leqslant \gamma_0/\psi(\alpha).$$

Otherwise, if $t \leqslant \alpha$ then

$$t\,|g_\alpha(t)|\,/\psi(t) = \sqrt{t}\,|g_\alpha(t)|\,\frac{\sqrt{t}}{\psi(t)} \leqslant \frac{\gamma_*}{\sqrt{\alpha}}\frac{\sqrt{\alpha}}{\psi(\alpha)},$$

and the proof is complete. □

Remark 5. A look at example 3, see (19), reveals that we used exactly the construction w_α corresponding to (A.2) and for spectral cutoff.

Remark 6. Note that for $\psi(t) = \sqrt{t},\ t > 0$, both elements v_α and w_α coincide. The proofs given below will distinguish between the two cases that ψ tends to zero slower than $t \mapsto \sqrt{t}$ (low-benchmark case) and faster than $t \mapsto \sqrt{t}$ (high-benchmark case).

The following is obvious:

$$\|x - \psi(H)v_\alpha\| = \|r_\alpha(\psi^2(H))x\| \tag{A.3}$$

and

$$\|x - \psi(H)w_\alpha\| = \|r_\alpha(H)x\|. \tag{A.4}$$

A.1. Low-benchmark case

Here we assume that the function $\psi^2(t)/t$ is non-increasing , and hence the element w_α from (A.2) is well-defined.

Lemma 6. *If the function $\psi^2(t)/t$ is non-increasing then*

$$d_\psi\left(2\max\{\gamma_0,\gamma_*\}\Phi_\psi^{-1}(\psi(\alpha))\right) \leqslant \|r_\alpha(H)x\|, \qquad 0 < \alpha \leqslant \bar{\alpha}. \tag{A.5}$$

Proof. Given α, we let R from $\psi(\alpha) = \Phi_\psi(R)$, and denote by v the minimizing element for $d_\psi(R)$. By using the element w_α from (A.2) we have by (A.4) that

$$d_\psi(\|w_\alpha\|) \leqslant \|r_\alpha(H)x\|,$$

such that it is enough to bound $\|w_\alpha\|$, appropriately:

$$\begin{aligned}\|w_\alpha\| &\leqslant \|g_\alpha(H)H\psi^{-1}(H)(x - \psi(H)v)\| + \|g_\alpha(H)H\psi^{-1}(H)\psi(H)v\| \\ &\leqslant \max\{\gamma_0,\gamma_*\}\frac{d_\psi(R)}{\psi(\alpha)} + \gamma_0 R \\ &\leqslant 2\max\{\gamma_0,\gamma_*\}R\max\left\{\frac{\Phi_\psi(R)}{\psi(\alpha)}, 1\right\} = 2\max\{\gamma_0,\gamma_*\}\,R.\end{aligned}$$

From this the proof can easily be completed. □

A.2. High-benchmark case

We turn to the the high-benchmark case, where ψ tends to zero faster than $t \mapsto \sqrt{t}$, and we start with the following observation.

Lemma 7. *Let g_α be any regularization. Then*

$$d_\psi((\gamma_0 + \gamma_*)\Phi_\psi^{-1}(\sqrt{\alpha})) \leqslant \|r_\alpha(\psi^2(H))x\|, \qquad 0 < \alpha \leqslant \bar{\alpha}. \tag{A.6}$$

If, in addition the function $t \mapsto \sqrt{t}$ is a qualification of the regularization then

$$\|r_\alpha(\psi^2(H))x\| \leqslant (\gamma + \gamma_*)d_\psi\left(\left(\Phi_\psi^{-1}(\sqrt{\alpha})\right)\right), \qquad 0 < \alpha \leqslant \bar{\alpha}. \tag{A.7}$$

Proof. We consider the family v_α from (A.1). In a first step we bound the norm of v_α. To this end let R be obtained from $\sqrt{\alpha} = \Phi_\psi(R)$, and denote v the element realizing the distance function $d_\psi(R)$. Then we have that

$$\begin{aligned}\|v_\alpha\| &\leqslant \|g_\alpha(\psi^2(H))\psi(H)(x - \psi(H)v)\| + \|g_\alpha(\psi^2(H))\psi^2(H)v\| \\ &\leqslant \frac{\gamma_*}{\sqrt{\alpha}} d_\psi(R) + \gamma_0 R = (\gamma_* + \gamma_0)R = (\gamma_* + \gamma_0)\Phi_\psi^{-1}(\sqrt{\alpha}),\end{aligned}$$

by the choice of R.

Since the distance function is decreasing we obtain that

$$d_\psi\left((\gamma_0 + \gamma_*)\Phi_\psi^{-1}(\sqrt{\alpha})\right) \leqslant d_\psi(\|v_\alpha\|) \leqslant \|x - \psi(H)v_\alpha\| = \|r_\alpha(\psi^2(H))x\|,$$

which yields (A.6).

Now, suppose that g_α has qualification as stated. Then we can argue, with element v as before, that

$$\begin{aligned}\|r_\alpha(\psi^2(H))x\| &\leqslant \|r_\alpha(\psi^2(H))(x - \psi(H)v)\| + \|r_\alpha(\psi^2(H))\psi(H)v\| \leqslant \gamma_1 d_\psi(R) + \gamma\sqrt{\alpha}R \\ &= (\gamma_1 + \gamma)d_\psi(R) \\ &= (\gamma_1 + \gamma)d_\psi\left(\left(\Phi_\psi^{-1}(\sqrt{\alpha})\right)\right).\end{aligned}$$

The proof is complete. □

In order to establish the required bound from theorem 1 we need to bound $\|r_\alpha(\psi^2(H))x\|$ from above. To this end we provide some estimate for regularization from a single function.

Lemma 8. *Suppose that ψ is an index function such that $\psi^2(t)/t$ is non-decreasing. If the function g gives rise for a linear regularization and if the accompanying function r has a non-increasing absolute value, and if $|r(1)| > 0$, then*

$$\left| r\left(\frac{\psi^2(t)}{\psi^2(\alpha)}\right)\right| \leqslant \frac{\gamma_1}{|r(1)|} r\left(\frac{t}{\alpha}\right), \qquad t, \alpha > 0. \tag{A.8}$$

Consequently,

$$\left|r_{\psi^2(\alpha)}(\psi^2(t))\right| \leqslant \frac{\gamma_1}{|r(1)|}\left|r_\alpha(t)\right|.$$

Proof. We consider two cases. If $t \leqslant \alpha$ then $0 < \psi^2(t)/\psi^2(\alpha) \leqslant 1$, and hence

$$\left| r\left(\frac{\psi^2(t)}{\psi^2(\alpha)}\right)\right| \leqslant |r(0)| = \gamma_1 \leqslant \frac{\gamma_1}{|r(1)|}|r(1)| \leqslant \frac{\gamma_1}{|r(1)|}\left| r\left(\frac{t}{\alpha}\right)\right|.$$

Otherwise, if $t > \alpha$ then $\frac{\psi^2(t)}{t} > \frac{\psi^2(\alpha)}{\alpha}$, and hence $\frac{\psi^2(t)}{\psi^2(\alpha)} > \frac{t}{\alpha}$, such that the monotonicity of $x \mapsto |r(x)|$ yields

$$\left| r\left(\frac{\psi^2(t)}{\psi^2(\alpha)}\right)\right| \leqslant \left| r\left(\frac{t}{\alpha}\right)\right|,$$

and the proof is complete, since $1 \leqslant \gamma_1/|r(1)|$. □

For Tikhonov regularization this is fulfilled with $\gamma_1/|r(1)| = 2$.

Remark 7. Note that in case of spectral cutoff we have the equality $\left|r_{\psi^2(\alpha)}(\psi^2(t))\right| = |r_\alpha(t)|$ for arbitrary index function ψ.

The bound for the high-benchmark case is now given in

Lemma 9. *Let g_α be a regularization which is obtained from a single function g with the non-increasing function r, $|r(1)| > 0$. Then*

$$d_\psi((\gamma_0 + \gamma_*)\Phi_\psi^{-1}(\psi(\alpha))) \leqslant \frac{\gamma_1}{|r(1)|}\|r_\alpha(H)x\|, \qquad \alpha > 0.$$

Proof. We shall apply lemma 7 with $\alpha := \psi^2(\alpha)$ and obtain that

$$d_\psi((\gamma_0 + \gamma_*)\Phi_\psi^{-1}(\sqrt{\alpha})) \leqslant \|r_{\psi^2(\alpha)}(\psi^2(H))x\|, \qquad \alpha > 0,$$

such that an application of lemma 8 allows us to complete the proof. □

Proof of theorem 1.

Lemmas 6 and 9 yield the first assertion in theorem 1. The upper bound was established in proposition 2, by noting that $\gamma_1 + \gamma \leqslant 2\max\{\gamma_1, \gamma\}$. □

References

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

[2] Boţ R I and Hofmann B 2010 An extension of the variational inequality approach for obtaining convergence rates in regularization of nonlinear ill-posed problems *J. Integral Eqns. Appl.* **22** 369–92

[3] 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

[4] 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

[5] Engl H W, Hanke M and Neubauer A 1996 *Regularization of Inverse Problems (Mathematics and its Applications* vol 375) (Dordrecht: Kluwer)

[6] Groetsch C W 1992 Spectral methods for linear inverse problems with unbounded operators *J. Approx. Theory* **70** 16–28

[7] Hanke M 1991 Accelerated Landweber iterations for the solution of ill-posed equations *Numer. Math.* **60** 341–73

[8] Hegland M and Hofmann B 2010 Errors of regularisation under range inclusions using variable Hilbert scales arXiv:1005.3883v1

[9] Hein T and Hofmann B 2009 Approximate source conditions for nonlinear ill-posed problems—chances and limitations *Inverse Problems* **25** 035003

[10] 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

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

[12] Hofmann B and Fleischer G 1999 Stability rates for linear ill-posed problems with compact and non-compact operators *Z. Anal. Anwendungen* **18** 267–86

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

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

[15] Hofmann B, Mathé P and von Weizsäcker H 2009 Regularization in Hilbert space under unbounded operators and general source conditions *Inverse Problems* **25** 115013

[16] Hofmann B, Schieck M and Wolfersdorf L v 2007 On the analysis of distance functions for linear ill-posed problems with an application to the integration operator in L^2 *J. Inverse Ill-Posed Probl.* **15** 83–98

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

[18] Kaltenbacher B and Hofmann B 2010 Convergence rates for the iteratively regularized Gauss–Newton method in Banach spaces *Inverse Problems* **26** 035007

[19] Krebs J, Louis A K and Wendland H 2009 Sobolev error estimates and *a priori* parameter selection for semi-discrete Tikhonov regularization *J. Inverse Ill-Posed Probl.* **17** 845–69

[20] Louis A K 1989 *Inverse und schlecht gestellte Probleme Teubner Studienbücher Mathematik (Teubner Mathematical Textbooks)* (Stuttgart: Teubner)

[21] Mathé P and Hofmann B 2008 Direct and inverse results in variable Hilbert scales *J. Approx. Theory* **154** 77–89

[22] Mathé P and Hofmann B 2008 How general are general source conditions? *Inverse Problems* **24** 015009

[23] Mathé P and Pereverzev S V 2003 Discretization strategy for linear ill-posed problems in variable Hilbert scales *Inverse Problems* **19** 1263–77

[24] Neubauer A 1994 On converse and saturation results for regularization methods *Beiträge zur angewandten Analysis und Informatik* (Aachen: Shaker) pp 262–70

[25] Neubauer A 1997 On converse and saturation results for Tikhonov regularization of linear ill-posed problems *SIAM J. Numer. Anal.* **34** 517–27

[26] Rudin W 1973 *Functional Analysis (McGraw-Hill Series in Higher Mathematics)* (New York: McGraw-Hill)

[27] Vaĭnikko G M and Veretennikov A Yu 1986 *Iteratsionnye protsedury v nekorrektnykh zadachakh* (Moscow: Nauka)

[28] Zălinescu C 2002 *Convex Analysis in General Vector Spaces* (River Edge, NJ: World Scientific)