CHARACTERIZATIONS OF VARIATIONAL SOURCE CONDITIONS, CONVERSE RESULTS, AND MAXISETS OF SPECTRAL REGULARIZATION METHODS

THORSTEN HOHAGE AND FREDERIC WEIDLING

Institute for Numerical and Applied Mathematics, University of Goettingen, Lotzestr. 16-18, 37083 Goettingen, Germany

Abstract. We describe a general strategy for the verification of variational source condition by formulating two sufficient criteria describing the smoothness of the solution and the degree of ill-posedness of the forward operator in terms of a family of subspaces. For linear deterministic inverse problems we show that variational source conditions are necessary and sufficient for convergence rates of spectral regularization methods, which are slower than the square root of the noise level. A similar result is shown for linear inverse problems with white noise. In many cases variational source conditions can be characterized by Besov spaces. This is discussed for a number of prominent inverse problems.

1. Introduction

This paper is concerned with inverse problems described by ill-posed operator equations in real Hilbert spaces $\mathbb{X}$ and $\mathbb{Y}$. Let $T : \mathbb{X} \to \mathbb{Y}$ an injective, bounded, linear operator and $f^\dagger \in \mathbb{X}$ the unknown solution to the inverse problem. We will study both a deterministic and a white noise noise model. In the first case measurement errors are described by a vector $\xi \in \mathbb{Y}$, and observed data are given by

$$g^{\text{obs}} = Tf^\dagger + \xi, \qquad \|\xi\| \leq \delta \tag{1}$$

for some deterministic noise level $\delta > 0$. In the second case measurement errors are described by a white noise process W on $\mathbb{Y}$, and observed data are given by

$$g^{\text{obs}} = Tf^\dagger + \varepsilon W \tag{2}$$

with a stochastic noise level $\varepsilon > 0$. Recall that a white noise process is characterized by the relations $\mathbf{E}[\langle W, y\rangle] = 0$ and $\mathbf{E}\left[\langle W, y_1\rangle\langle W, y_2\rangle\right] = \langle y_1, y_2\rangle$ for all $y, y_1, y_2 \in \mathbb{Y}$.

Regularization theory is concerned with error estimates for approximate reconstruction methods (regularization methods) for $f^\dagger$ given data g^{obs} described by (1) or (2). It is well-known that for ill-posed problems uniform error bounds necessarily require further assumptions on the solution $f^\dagger$ (see [6, Prop. 3.11]). Such conditions are usually called

E-mail address: t.hohage@math.uni-goettingen.de,f.weidling@math.uni-goettingen.de.

source conditions. Over the last years, starting with [9] it has become increasingly popular to formulate such conditions in the form of variational inequalities

$$2\left\langle f^\dagger, f^\dagger - f\right\rangle_{\mathbb{X}} \leq \frac{1}{2}\left\|f - f^\dagger\right\|_{\mathbb{X}}^2 + \psi\left(\left\|T(f) - T(f^\dagger)\right\|_{\mathbb{Y}}^2\right) \qquad \text{for all } f \in \mathbb{X} \tag{3}$$

with an index function ψ. (A function $\psi : [0,\infty) \to [0,\infty)$ is called an *index function* if it is continuous, strictly monotonically increasing, and $\psi(0) = 0$.) Advantages of these *variational source conditions (VSC)* over classical source conditions of the form $f^\dagger = \varphi(T^*T)w$ with an index function φ and $w \in \mathbb{X}$ include extensions to Banach spaces, general penalty and data fidelity functionals, treatment of nonlinear operators without the need of a derivative of T and restrictive assumptions relating T' and T, as well as simpler proofs. As a disadvantage we mention that (3) cannot be used to describe high order rates of convergence since it is easy to see that it cannot hold true for $f^\dagger \neq 0$ if $\lim_{x\to 0} \psi(x^2)/x = 0$. This excludes in particular the case $\psi(t) = t^\nu$ with $\nu > 1/2$.

In this paper we will address the following two related main questions:

- What are verifiable sufficient (and possibly even necessary) conditions such that the VSC (3) holds true?
- What are necessary and sufficient conditions on $f^\dagger$ for a given rate of convergence of a given regularization method as the noise level δ or ε tends to 0?

Let us now discuss known results from the literature and the contributions of this paper for both of these questions. Concerning the first question, verifiable sufficient conditions for (3) have mainly been given via spectral source conditions so far, see Appendix A for more details. In [11, 23] we have recently derived sufficient conditions for (3) in the form of bounds on Sobolev norms of $f^\dagger$ for nonlinear inverse medium scattering problems. Here we formulate in Theorem 2.1 two criteria which capture the main strategy of the proofs in [11, 23]. In the following we will apply them to linear inverse problems. These criteria describe the two main factors influencing rates of convergence: Smoothness of the solution and ill-posedness of the inverse of the forward operator. Here both criteria are formulated in terms of a sequence of approximating subspaces in $\mathbb{X}$. If these spaces are chosen as eigenspaces of T^*T, we obtain an equivalent characterization of the VSC (3) in terms of the rate of decay of the spectral distribution function of $f^\dagger$. The latter criterion has been introduced by Neubauer [17] and shown to be necessary and sufficient for Hölder rates of convergence. A characterization of the VSC (3) for $\psi(t) = t^\nu$, $\nu \in (0, 1/2]$ has previously been shown by Andreev et al. [4] using a different technique which seems to be limited to this special case.

Let us now discuss the second question. In deterministic regularization theory theorems on the necessity of conditions for a given rate of convergence (which have already shown to be sufficient) are known as *converse results*. In statistics the maximal set of $f^\dagger$ for which a given estimator achieves a given desired rate of convergence is called a *maxiset*. Converse results for Hölder rates of Tikhonov regularization have been established by Neubauer [17]. Andreev [3] has proven converse results for Hölder rates of (generalized) Tikhonov regularization and Landweber iteration in terms of K-interpolation spaces with fine index $q = \infty$ between $\mathbb{X}$ and $(T^*T)^\nu(\mathbb{X})$ equipped with the image space norm. Flemming, Hofmann &

Mathé [8] have derived converse results for general convergence rates of the bias of general spectral regularization methods in terms of approximate source conditions. Albani et al. [1] proved converse results for general deterministic rates and spectral regularization method, but additional assumptions had to be imposed, which are not always obvious to interpret. Here we will prove converse results without such assumptions. As a byproduct of our analysis we show the equivalence of weak and strong quasioptimality of a posteriori parameter choice rules in many cases (see [20]). Together with our answer to the first question we also obtain converse results in terms of VSCs (3) for concave ψ. Moreover, we will show for inverse problems for which the forward operator satisfies $T^*T = \Lambda(-\Delta)$ for some Laplace-Beltrami operator Δ, that VSCs for certain index functions ψ are satisfied if and only if $f^\dagger$ belongs to a Besov space $B^s_{2,\infty}$. This holds true in particular for the backward and sideways heat equation, and the inverse gradiometry problem.

In statistics maxisets of wavelet methods for the estimation of the density of i.i.d. random variables have been characterized as Besov spaces by Kerkyacharian & Picard [12]. They consider not only L^2, but also other L^p norms as loss functions. Maxisets of thresholding and more general wavelet estimators have been characterized by the same authors in [13, 14], and their results have been generalized by Rivoirand [21] to some linear estimators in the sequence space model of inverse problems. These latter references show in particular that under certain circumstances nonlinear thresholding methods have larger maxisets than linear methods for given polynomial rates. Here we show under fairly general assumptions that VSCs characterize maxisets of spectral regularization methods for linear inverse problems with white noise.

The plan of this paper is as follows: In the following section we formulate and prove the theorem describing our general strategy for the verification of VSCs. In sections 3–5 we derive converse results for the bias, rates of convergence with deterministic noise, and rates of convergence with white noise, respectively. In Section 6 we introduce a class of inverse problems for which maxisets of linear spectral regularization methods are given by Besov spaces $B^s_{2,\infty}$. Finally, in Section 7 we apply our theoretical results to a number of well-known inverse problem before we end this paper with some conclusions. In an appendix we show how the general strategy in Section 2 can be applied to verify a VSC given a spectral source condition for linear problems.

2. A general strategy for verifying variational source conditions

In this section we formulate sufficient conditions for VSCs in terms of arbitrary families of subspaces. In the rest of this paper these will always be chosen as invariant subspaces of T^*T, but in principle the choice is arbitrary. To allow also polynomial, trigonometric, wavelet and other subspaces, which may be relevant in more general situation (see e.g. [11]), we will parametrize the spaces by a general index set $\mathcal{J}$.

Theorem 2.1. *Suppose there exists a family of orthogonal projections $P_r \in L(\mathbb{X})$ indexed by a parameter r in some index set $\mathcal{J}$ such that for some functions $\kappa, \sigma : \mathcal{J} \to (0,\infty)$ and some $C \geq 0$ the following conditions hold true for all $r \in \mathcal{J}$:*

$$\|f^\dagger - P_r f^\dagger\|_{\mathbb{X}} \leq \kappa(r) \tag{4}$$

$$\left\langle f^\dagger, P_r(f^\dagger - f)\right\rangle_{\mathbb{X}} \leq \sigma(r)\|T(f^\dagger) - T(f)\| + C\kappa(r)\|f^\dagger - f\| \quad \textit{for all } f \in D(T) \textit{ with } \|f - f^\dagger\| \leq 4\|f^\dagger\|. \tag{5}$$

Then $f^\dagger$ *satisfies the VSC* (3) *with*

$$\psi(t) := 2\inf_{r\in\mathcal{J}}\left[(C+1)^2\kappa(r)^2 + \sigma(r)\sqrt{t}\right].$$

Assumption (4) may be seen as an estimate of the approximation quality of the approximating spaces $P_r\mathbb{X}$, and assumption (5) may be seen as a kind of stability estimate for T on these spaces. If $f^\dagger$ belongs to some smooth subspace of $\mathbb{X}$, the stability estimate may be taken with respect to the norm of the dual space. However, (5) is not exactly a stability estimate for the restriction of T to $P_r\mathbb{X}$ since by do not bound by $\|T(P_rf) - T(P_rf^\dagger)\|$, but $\|T(f) - T(f^\dagger)\|$. On the other hand the additional term $C\kappa(r)\|f - f^\dagger\|$ may help. The case $C > 0$ and the restriction to $f \in \mathbb{X}$ with $\|f - f^\dagger\| \leq 4\|f^\dagger\|$ are also crucial for nonlinear inverse problems.

Proof of Theorem 2.1. If $\|f - f^\dagger\| > 4\|f^\dagger\|$ we have

$$2\left\langle f^\dagger, f^\dagger - f\right\rangle \leq 2\|f^\dagger\|\,\|f^\dagger - f\| \leq \frac{1}{2}\|f^\dagger - f\|^2, \tag{6}$$

so (3) holds true. Otherwise we can apply (5) and (4) and the basic inequality $2ab \leq 2a^2 + \frac{1}{2}b^2$ to obtain

$$\begin{aligned}
2\left\langle f^\dagger, (f^\dagger - f)\right\rangle &= 2\left\langle f^\dagger, P_r(f^\dagger - f)\right\rangle + 2\left\langle (I - P_r)f^\dagger, f^\dagger - f\right\rangle \\
&\leq 2\sigma(r)\|T(f) - T(f^\dagger)\| + 2(C+1)\kappa(r)\|f^\dagger - f\| \\
&\leq 2\sigma(r)\|T(f) - T(f^\dagger)\| + 2(C+1)^2\kappa(r)^2 + \frac{1}{2}\|f^\dagger - f\|^2
\end{aligned}$$

for all $r \in \mathcal{J}$. Taking the infimum over of the right hand side over $r \in \mathcal{J}$ with $t = \|T(f) - T(f^\dagger)\|^2$ yields (3). □

3. Converse results for the bias

For $\lambda \geq 0$ we define the spectral projections

$$E_\lambda^{T^*T} := \chi_{[0,\lambda]}(T^*T). \tag{7}$$

The function $\lambda \mapsto \|E_\lambda^{T^*T} f\|$ is called the *spectral distribution function* of $f \in \mathbb{X}$. For an index function κ we define a subspace $\mathbb{X}_\kappa^T \subset \mathbb{X}$ via a weighted supremum norm of the spectral distribution function with weight $1/\kappa$:

$$\mathbb{X}_\kappa^T := \left\{f \in \mathbb{X}\colon \|f\|_{\mathbb{X}_\kappa^T} < \infty\right\}, \qquad \|f\|_{\mathbb{X}_\kappa^T} := \sup_{\lambda>0}\frac{1}{\kappa(\lambda)}\left\|E_\lambda^{T^*T} f\right\|_{\mathbb{X}} \tag{8}$$

It is not difficult to see that $\mathbb{X}_\kappa^T$ is a Banach space. Note that the unit ball in $\mathbb{X}_\kappa^T$ contains all functions satisfying (4) with $P_r = I - E_r^{T^*T}$. For the remainder of this and the following two sections we will suppress the superscript T^*T to simplify the notation. However, we will need it in Section 6 to deal with spectral distribution functions w.r.t. several operators simultaneously.

Theorem 3.1. *Let $\kappa : [0,\infty) \to [0,\infty)$ be an index function such that $t \mapsto \kappa(t)^2/t^{1-\mu}$ is decreasing for some $\mu \in (0,1)$ and $\kappa \cdot \kappa$ is concave. Moreover, we associate with each such κ an index function ψ_κ by*

$$\psi_\kappa(t) := \kappa\left(\Theta_\kappa^{-1}(\sqrt{t})\right)^2, \qquad \Theta_\kappa(\lambda) := \sqrt{\lambda}\kappa(\lambda). \tag{9}$$

Then the following statements for $f^\dagger \in \mathbb{X}$ are equivalent:

(i) *$f^\dagger$ satisfies a VSC with $\psi(t) = A\psi_\kappa(t)$ for some $A > 0$.*

(ii) *$\|f^\dagger\|_{\mathbb{X}_\kappa^T} < \infty$.*

More precisely, (i) implies $\|E_\lambda f^\dagger\| \le \sqrt{\frac{2A}{3}}\kappa\left(\frac{2A}{3}\lambda\right) \le \sqrt{\frac{2A}{3}}\max(1,\sqrt{\frac{2A}{3}})\kappa(\lambda)$. Vice versa, if κ is normalized such that $\|f^\dagger\|_{\mathbb{X}_\kappa^T} = 1$, then (i) holds true with $A = 2(1+\mu^{-1}) + 2\kappa(\|T\|^2)\sup_{t\in(0,4\|T\|\|f^\dagger\|]}\sqrt{t}/\psi_\kappa(t)$, and A is finite.

Proof. (i) $\Rightarrow$ (ii): First note that

$$\psi_\kappa^{-1}(\xi) = \xi \cdot (\kappa\cdot\kappa)^{-1}(\xi) \tag{10}$$

since

$$\psi_\kappa(t)\cdot(\kappa\cdot\kappa)^{-1}(\psi_\kappa(t)) = \kappa(\Theta_\kappa^{-1}(\sqrt{t}))^2\cdot\Theta_\kappa^{-1}(\sqrt{t}) = \Theta_\kappa(\Theta_\kappa^{-1}(\sqrt{t}))^2 = \sqrt{t}^2 = t.$$

Choosing $f = (I - E_\lambda)f^\dagger$ in (3) yields

$$2\|E_\lambda f^\dagger\|^2 = 2\left\langle f^\dagger, E_\lambda f^\dagger\right\rangle \le \frac{1}{2}\|E_\lambda f^\dagger\|^2 + A\psi_\kappa\left(\|TE_\lambda f^\dagger\|^2\right).$$

As

$$\|TE_\lambda f^\dagger\|^2 = \int_0^\lambda \lambda \,\mathrm{d}\|E_\lambda f^\dagger\|^2 \le \lambda\int_0^\lambda \mathrm{d}\|E_{\tilde\lambda} f^\dagger\|^2 = \lambda\|E_{\tilde\lambda} f^\dagger\|^2,$$

the spectral distribution function $\tilde\kappa(\lambda) := \|E_\lambda f^\dagger\|$ of $f^\dagger$ satisfies the inequality $\frac{3}{2}\tilde\kappa(\lambda)^2 \le A\psi_\kappa\left(\lambda\tilde\kappa(\lambda)^2\right)$. Hence, setting $\widetilde{\psi}_\kappa(t) := \psi_\kappa(t)/t$ we have

$$\frac{3}{2A\lambda} \le \frac{\psi_\kappa\left(\lambda\tilde\kappa(\lambda)^2\right)}{\lambda\tilde\kappa(\lambda)^2} = \widetilde{\psi}_\kappa(\lambda\tilde\kappa(\lambda)^2).$$

As $\kappa\cdot\kappa$ is assumed to be a concave index function, the inverse function $(\kappa\cdot\kappa)^{-1}$ is a convex index function. This implies that $\xi \mapsto \xi\cdot(\kappa\cdot\kappa)^{-1}(\xi)$ is a convex index function as well, and from (10) we see that ψ_κ is concave. This in turn implies that $\widetilde{\psi}_\kappa$ is monotonically decreasing, so we obtain

$$\widetilde{\psi}_\kappa^{-1}\left(\frac{3}{2A\lambda}\right) \ge \lambda\tilde\kappa(\lambda)^2$$

or with $\beta = \frac{2A\lambda}{3}$

$$\tilde\kappa\left(\frac{3}{2A}\beta\right) \le \sqrt{\frac{2A}{3}}\sqrt{\frac{1}{\beta}\widetilde{\psi}_\kappa^{-1}\left(\frac{1}{\beta}\right)} \tag{11}$$

for all $\beta > 0$. Setting $\xi = \kappa(\beta)^2$ in (10), we obtain $\psi_\kappa^{-1}(\kappa(\beta)^2) = \beta\kappa(\beta)^2$ or $\kappa(\beta)^2 = \psi_\kappa(\beta\kappa(\beta)^2)$. This is equivalent to $\frac{1}{\beta} = \frac{\psi_\kappa(\beta\kappa(\beta)^2)}{\beta\kappa(\beta)^2}$ and to $\frac{1}{\beta}\widetilde{\psi}_\kappa^{-1}\left(\frac{1}{\beta}\right) = \kappa^2(\beta)$. Plugging this into (11) shows that $\tilde{\kappa}(\lambda) \leq \sqrt{\frac{2A}{3}}\kappa\left(\frac{2A}{3}\lambda\right)$ for all $\lambda > 0$. Due to the concavity of the index function $\kappa \cdot \kappa$ we have $\kappa\left(\frac{2A}{3}\lambda\right)^2 \leq \kappa(\max(1, \frac{2A}{3})\lambda)^2 \leq \max(1, \frac{2A}{3})\kappa(\lambda)^2$.

(ii) $\Rightarrow$ *(i)*: Suppose that $\|f^\dagger\|_{\mathbb{X}_\kappa^T} = 1$, i.e. $\|E_\lambda f^\dagger\| \leq \kappa(\lambda)$ for all $\lambda > 0$. In a first step we show that

$$\|(T(I - E_\lambda))^\dagger f^\dagger\|^2 = \int_\lambda^{\|T^*T\|} \frac{1}{t}\,\mathrm{d}\|E_t f^\dagger\|^2 \leq \frac{1}{\mu}\frac{\kappa(\lambda)^2}{\lambda} + \|f^\dagger\|^2 \tag{12}$$

for all $\lambda > 0$. Here $(T(I - E_\lambda))^\dagger$ denotes the Moore-Penrose inverse of $T(I - E_\lambda)$. By partial integration we have

$$\int_\lambda^{\|T^*T\|} \frac{1}{t}\,\mathrm{d}\|E_t f^\dagger\|^2 = \|f^\dagger\|^2 - \|E_\lambda f^\dagger\|^2 + \int_\lambda^{\|T^*T\|} \frac{\|E_t f^\dagger\|^2}{t^2}\,\mathrm{d}t.$$

Now we use the assumption $\|E_t f^\dagger\| \leq \kappa(t)$ and the monotonicity of $\kappa(t)^2/t^{1-\mu}$ to obtain

$$\begin{aligned}\int_\lambda^{\|T^*T\|} \frac{\|E_t f^\dagger\|^2}{t^2}\,\mathrm{d}t &\leq \int_\lambda^{\|T^*T\|} \frac{\kappa(t)^2}{t^{1-\mu}}\frac{1}{t^{1+\mu}}\,\mathrm{d}t \leq \frac{\kappa(\lambda)^2}{\lambda^{1-\mu}}\int_\lambda^{\|T^*T\|} \frac{1}{t^{1+\mu}}\,\mathrm{d}t \\ &= \frac{\kappa(\lambda)^2}{\mu\lambda}\left(1 - \frac{\lambda^\mu}{\|T^*T\|^\mu}\right) \leq \frac{\kappa(\lambda)^2}{\mu\lambda}.\end{aligned}$$

In a second step we can now use (12) to verify assumption (5) in Theorem 2.1:

$$\begin{aligned}\left\langle f^\dagger, (I - E_\lambda)(f^\dagger - f)\right\rangle &= \left\langle (T(I - E_\lambda))^\dagger f^\dagger, T(I - E_\lambda)(f^\dagger - f)\right\rangle \\ &\leq \|T(I - E_\lambda))^\dagger f^\dagger\|\,\|T(I - E_\lambda)(f^\dagger - f)\| \leq \left(\frac{\kappa(\lambda)}{\sqrt{\mu\lambda}} + \|f^\dagger\|\right)\|Tf^\dagger - Tf\|,\end{aligned}$$

i.e. $\sigma(\lambda) = \frac{\kappa(\lambda)}{\mu\lambda} + \|f^\dagger\|$. Hence, by Theorem 2.1 (3) holds true with

$$\psi(t) = 2\inf_{\lambda > 0}\left[\kappa(\lambda)^2 + \left(\frac{\kappa(\lambda)}{\sqrt{\mu\lambda}} + \|f^\dagger\|\right)\sqrt{t}\right] \leq 2\left(1 + \frac{1}{\sqrt{\mu}}\right)\psi_\kappa(t) + 2\|f^\dagger\|\sqrt{t}$$

where we have chosen $\lambda = \Theta_\kappa^{-1}(\sqrt{t})$, i.e. $\sqrt{\lambda}\kappa(\lambda) = \sqrt{t}$. This implies $\frac{\kappa(\lambda)}{\sqrt{\lambda}}\sqrt{t} = \kappa^2(\lambda) = \psi_\kappa(t)$. It remains to bound $\sqrt{t}$ in terms of $\psi_\kappa(t)$. Note from (6) that we only need to show the variational source condition for $\|f^\dagger - f\| \leq 4\|f^\dagger\|$ in order to prove it everywhere. Hence it is enough to bound $\sqrt{t}$ by $\psi_\kappa(t)$ for $\sqrt{t} = \|Tf^\dagger - Tf\| \leq 4\|T\|\|f^\dagger\|$. We have

$$\|f^\dagger\|\sqrt{t} \leq \kappa(\|T\|^2)\psi_\kappa(t) \sup_{\tau \in (0, 4\|T\|\|f^\dagger\|]} \frac{\sqrt{\tau}}{\psi_\kappa(\tau)}.$$

To see that this is finite and even $\lim_{\tau\to 0}\sqrt{\tau}/\psi_\kappa(\tau) = 0$, we substitute $\delta = \Theta_\kappa^{-1}(\sqrt{\tau})$:

$$\lim_{\tau\to 0}\frac{\sqrt{\tau}}{\psi_\kappa(\tau)} = \lim_{\tau\to 0}\frac{\sqrt{\tau}}{\kappa(\Theta_\kappa^{-1}(\sqrt{\tau}))^2} = \lim_{\delta\to 0}\frac{\Theta_\kappa(\delta)}{\kappa(\delta)^2} \leq \lim_{\delta\to 0}\delta^{\mu/2}\sqrt{\frac{\delta^{1-\mu}}{\kappa^2(\delta)}} = 0. \qquad \square$$

We now consider spectral regularization methods of the form

$$\widehat{f}_\alpha := R_\alpha g^{\mathrm{obs}} \qquad \text{with} \qquad R_\alpha := q_\alpha(T^*T)T^* \tag{13}$$

and impose the following assumptions:

Assumption 3.2. With $r_\alpha(\lambda) := 1 - \lambda q_\alpha(\lambda)$ assume that there are constants $C_1 > 0$, $\overline{\alpha} \in (0, \infty]$, and $0 < C_2 \leq C_3 < 1$ such that

(i) $|q_\alpha(\lambda)| \leq \frac{C_1}{\alpha}$ for all $\lambda \in [0, \|T^*T\|]$,
(ii) $\lambda \mapsto r_\alpha(\lambda)$ is decreasing for all $\alpha > 0$ and $r_\alpha(\lambda) \geq 0$,
(iii) $\alpha \mapsto r_\alpha(\lambda)$ is increasing for all $\lambda \in [0, \|T^*T\|]$,
(iv) $C_2 \leq r_\alpha(\alpha) \leq C_3$ for all $0 < \alpha \leq \overline{\alpha}$.

As $r_\alpha(0) = 1 - 0q_\alpha(0) = 1$, assumption (ii) implies

$$0 \leq r_\alpha(\lambda) \leq 1 \tag{14}$$

for all $\alpha > 0$ and $\lambda \geq 0$. Below we will use the following notations for $x, y \in \mathbb{R}$:

$$x \vee y := \max(x, y), \qquad x \wedge y := \min(x, y)$$

Assumption 3.2 is satisfied in particular for the following methods. Unless stated otherwise we choose $\overline{\alpha} = \infty$. For a more detailed discussion of these methods we refer to [6].

- *Tikhonov regularization:* Here $q_\alpha(\lambda) = (\alpha + \lambda)^{-1}$ and $r_\alpha(\lambda) = \alpha/(\alpha + \lambda)$. We have $C_1 = 1$ and $C_2 = C_3 = \frac{1}{2}$.
- *Showalter's method:* Here $r_\alpha(\lambda) = \exp(-\lambda/\alpha)$, $C_1 = 1$ and $C_2 = C_3 = \exp(-1)$.
- *Landweber iteration:* For $\alpha > 0$ let $k_\alpha := \min\{n \in \mathbb{N}_0 : n + 1 > 1/\alpha\}$ be the number of iterations. Then $r_\alpha(\lambda) = (1 - \mu\lambda)^{k_\alpha}$ and $q_\alpha(\lambda) = \sum_{j=0}^{k_\alpha - 1}(1 - \mu\lambda)^j$ where $0 < \mu \leq \|T^*T\|^{-1}$ is the step length parameter. We have $C_1 = 1$. For $\alpha = 1/(n + \epsilon)$ with $\epsilon \in [0, 1)$ we have $k_\alpha = n$, therefore $r_\alpha(\alpha) \geq (1 - \mu/n)^n$ which by the inequality of arithmetic and geometric means is monotonically increaing in $n = k_\alpha$ for $k_\alpha > \mu$, hence we choose $\overline{\alpha} < \|T^*T\| \wedge 1$ and get $C_2 = (1 - \mu/k_{\overline{\alpha}})^{k_{\overline{\alpha}}}$ and $C_3 = \lim_{n\to\infty}(1 - \mu/(n + 1))^n = \exp(-\mu)$.
- *k-times iterated Tikhonov regularization:* This is described by $r_\alpha(\lambda) = \alpha^k/(\alpha+\lambda)^k$. We have $C_1 = k$ and $C_2 = C_3 = 2^{-k}$.
- *Lardy's method:* Here $r_\alpha(\lambda) = \beta^{k_\alpha}/(\beta+\lambda)^{k_\alpha}$ where $\beta > 0$ is fixed and the iteration number k_α and $C_1 = 1$. Choosing $\overline{\alpha} := 1 \wedge \beta$ we have $C_3 = \exp(-1/2\beta))$. Choosing α as for Landweber we see that $r_\alpha(\alpha) \geq (1+1/(\beta n))^{-n}$, therefore $C_2 = \exp(-1/\beta)$ since $(1 + 1/(\beta n))^n \to \exp(1/\beta)$ is monotonically increasing in n as argued above.
- *modified spectral cutoff:* Here $r_\alpha(\lambda) = (1 - \lambda/2\alpha) \vee 0$, $q_\alpha(\lambda) = 1/\lambda \wedge 1/2\alpha$, and $C_1 = C_2 = C_3 = 1/2$.

Note that Assumption 3.2(iv) is violated for standard spectral cutoff (or truncated SVD), i.e. $r_\alpha(\lambda) = 1$ if $\alpha \leq \lambda$ and $r_\alpha(\lambda) = 0$ else.

Theorem 3.3 ([1, Prop. 2.3])**.** *Assume that a spectral regularization method satisfies Assumption 3.2. Moreover, assume that for the index function κ there exists $A > 0$ and $\nu > 1$*

such that

$$r_\alpha(\lambda)\kappa(\lambda)^\nu \leq B\kappa(\alpha)^\nu \tag{15}$$

for all $\alpha, \lambda > 0$ *(i.e. the qualification of the regularization method covers* κ^ν *in the terminology of* [16]*). Then the following statements for* $f^\dagger \in \mathbb{X}$ *are equivalent:*

(i) $\|f^\dagger\|_{\mathbb{X}^T_\kappa} < \infty$.

(ii) $A := \sup_{0<\alpha\leq\overline{\alpha}} \frac{1}{\kappa(\alpha)}\|r_\alpha(T^*T)f^\dagger\| < \infty$*, i.e. the bias for* $f^\dagger$ *is of order* $\mathcal{O}(\kappa(\alpha))$*.*

More precisely,

$$\|f^\dagger\|_{\mathbb{X}^T_\kappa} \leq \frac{A}{C_2} \vee \frac{\|f^\dagger\|}{\kappa(\overline{\alpha})} \quad \textit{and} \quad A^2 \leq \frac{B\|f^\dagger\|^2}{\kappa(\|T\|^2)} + \|f^\dagger\|^2_{\mathbb{X}^T_\kappa}\left(1 + \frac{B^{1/\nu}\nu C_3^{(\nu-1)/\nu}}{\nu - 1}\right).$$

Proof. The more difficult implication (i) $\Rightarrow$ (ii) has been proved in [1, Prop. 2.3]. Since we have slightly different assumptions we give the proof of the implication (ii) $\Rightarrow$ (i). For $0 < \alpha \leq \overline{\alpha}$ we have

$$\|E_\alpha f^\dagger\| \leq \frac{1}{r_\alpha(\alpha)}\|r_\alpha(T^*T)f^\dagger\| \leq \frac{A}{C_2}\kappa(\alpha).$$

Otherwise, if $\alpha > \overline{\alpha}$, we have $\|E_\alpha f^\dagger\|/\kappa(\alpha) \leq \|f^\dagger\|/\kappa(\overline{\alpha})$. □

Recall that the largest number $\mu_0 > 0$ for which (15) holds true for $\kappa(\alpha)^\nu = \alpha^{\mu_0}$ is called the *classical qualification* of the regularization method. We have $\mu_0 = k$ for k-times iterated Tikhonov regularization and $\mu = \infty$ for Showalter's method, Lardy's methods, Landweber iteration, and modified spectral cutoff ([6]).

4. Converse results for deterministic noise

This section discusses regularization methods for the deterministic noise model (1).

Theorem 4.1. *Assume that a spectral regularization method satisfies Assumption 3.2. Moreover, let* κ *be an index function for which there exists* $p \geq 1$ *such that*

$$\kappa(r\alpha) \leq r^p\kappa(\alpha) \tag{16}$$

for all $\alpha > 0$ *and* $r \geq 1$ *(i.e.* κ *does not grow faster than polynomially). Then the following statements are equivalent for* $f^\dagger \in \mathbb{X}$*:*

(i) $A := \sup_{0<\alpha\leq\overline{\alpha}} \frac{1}{\kappa(\alpha)^2}\|r_\alpha(T^*T)f^\dagger\|^2 < \infty$.

(ii) $B := \sup_{0<\delta\leq\Theta_\kappa(\overline{\alpha})} \frac{1}{\psi_\kappa(\delta^2)} \inf_{0<\alpha\leq\overline{\alpha}} \sup_{\|\xi\|\leq\delta} \|R_\alpha(Tf^\dagger + \xi) - f^\dagger\|^2 < \infty$.

More precisely,

$$B \leq 2(A + C_1) \qquad \textit{and} \qquad A \leq B\left(\frac{4B^2}{(1-C_3)^4} \vee 1\right)^p \vee \|f^\dagger\|^2\kappa\left(\frac{\overline{\alpha}(1-C_3)^2}{2B}\right)^{-2}.$$

Proof. (i) $\Rightarrow$ *(ii):* From the standard estimate

$$\|R_\alpha\|^2 = \|R_\alpha^* R_\alpha\| = \|q_\alpha(T^*T)^2 T^*T\| \leq \|q_\alpha\|_\infty \|1 - r_\alpha\|_\infty \leq \frac{C_1}{\alpha} \tag{17}$$

using Assumption 3.2(i) and (14). Hence we have

$$\|R_\alpha(Tf^\dagger + \xi) - f^\dagger\|^2 \le \left(\|r_\alpha(T^*T)f^\dagger\| + \|R_\alpha\|\delta\right)^2 \le 2A\kappa(\alpha)^2 + 2\frac{C_1\delta^2}{\alpha}$$

for all $\|\xi\| \le \delta$ and $0 < \alpha \le \overline{\alpha}$. Choosing $\alpha = \Theta_\kappa^{-1}(\delta)$ and using $\sqrt{\Theta_\kappa^{-1}(\delta)}\kappa(\Theta_\kappa^{-1}(\delta)) = \Theta_\kappa(\Theta_\kappa^{-1}(\delta)) = \delta$, i.e. $\delta^2/\Theta_\kappa^{-1}(\delta) = \psi_\kappa(\delta^2)$, we obtain

$$\sup_{\|\xi\|\le\delta} \|R_\alpha(Tf^\dagger + \xi) - f^\dagger\|^2 \le (2A + 2C_1)\psi_\kappa(\delta^2).$$

(ii) $\Rightarrow$ *(i)*: Expanding

$$\|R_\alpha(Tf^\dagger+\xi)-f^\dagger\|^2 = \|r_\alpha(T^*T)f^\dagger+R_\alpha\xi\|^2 = \|r_\alpha(T^*T)f^\dagger\|^2+2\left\langle r_\alpha(T^*T)f^\dagger, R_\alpha\xi\right\rangle+\|R_\alpha\xi\|^2,$$

we see that only the middle of the three terms on the right hand side is affected by a sign change of ξ. Therefore, to bound the supremum over ξ from below, we may assume that the middle term is positive and neglect it to obtain

$$\sup_{\|\xi\|\le\delta} \|R_\alpha(Tf^\dagger+\xi)-f^\dagger\|^2 \ge \|r_\alpha(T^*T)f^\dagger\|^2+ \sup_{\|\xi\|\le\delta} \|R_\alpha\xi\|^2 = \|r_\alpha(T^*T)f^\dagger\|^2+\|R_\alpha\|^2\delta^2. \tag{18}$$

From the equality in (17) and the isometry of the functional calculus together with the last point in Assumption 3.2 we obtain

$$\|R_\alpha\|^2 = \sup_{\lambda\ge 0} \lambda|q_\alpha(\lambda)|^2 = \sup_{\lambda\ge 0} \frac{(1-r_\alpha(\lambda))^2}{\lambda} \ge \frac{(1-r_\alpha(\alpha))^2}{\alpha} \ge \frac{(1-C_3)^2}{\alpha}$$

if $0 < \alpha \le \overline{\alpha}$. By Assumption 3.2(iii) the first term on the right hand side of (18) is increasing in α whereas the second term is decreasing. Therefore, using the choice $\alpha^*(\delta) = \Theta_\kappa^{-1}(\delta)(1 - C_3)^2/(2B)$ from the first part of the proof, for which both terms are of the same order, we obtain the lower bound

$$\begin{aligned} B\psi_\kappa(\delta^2) &\ge \inf_{\alpha>0} \sup_{\|\xi\|\le\delta} \|R_\alpha(Tf^\dagger + \xi) - f^\dagger\|^2 \ge \|r_{\alpha^*(\delta)}(T^*T)f^\dagger\|^2 \wedge \|R_{\alpha^*(\delta)}\|^2\delta^2 \\ &\ge \|r_{\alpha^*(\delta)}(T^*T)f^\dagger\|^2 \wedge \frac{(1-C_3)^2}{\alpha^*(\delta)}\delta^2 = \|r_{\alpha^*(\delta)}(T^*T)f^\dagger\|^2 \wedge 2B\psi_\kappa(\delta^2) \end{aligned}$$

for $\overline{\alpha} \ge \alpha^*(\delta) > 0$. As $\|r_{\alpha^*(\delta)}(T^*T)f^\dagger\|^2 \ge 2B\psi_\kappa(\delta^2)$ would lead to a contradiction, the minimum is attained at the first argument, and we have $\|r_{\alpha^*(\delta)}(T^*T)f^\dagger\|^2 \le B\psi_\kappa(\delta^2)$. As $\delta = \Theta_\kappa(\frac{2B}{(1-C_3)^2}\alpha^*(\delta))$ and $\psi_\kappa(t) = (\kappa \circ \Theta_\kappa^{-1}(\sqrt{t}))^2$, we obtain

$$\|r_{\alpha^*}(T^*T)f^\dagger\|^2 \le B\kappa\left(\frac{2B}{(1-C_3)^2}\alpha^*\right)^2 \le B\left(\frac{4B^2}{(1-C_3)^4} \vee 1\right)^p \kappa(\alpha^*)^2$$

for all $0 < \alpha^* \le \overline{\alpha}(1 - C_3)^2/(2B)$. If $\overline{\alpha}(1 - C_3)^2/(2B) < \alpha \le \overline{\alpha}$ we can bound

$$\kappa(\alpha)^{-2}\|r_\alpha(T^*T)f^\dagger\|^2 \le \kappa\left(\frac{\overline{\alpha}(1-C_3)^2}{2B}\right)^{-2} \|f^\dagger\|^2$$

finishing the proof. ∎

We point out that in comparison to similar results by Neubauer [17, Thm. 2.6] and Albani et al. [1, Prop. 3.3] we have interchanged the order of the supremum over the noise vector ξ and the infimum over the regularization parameter α. Since obviously

$$\sup_{\|\xi\|\leq\delta}\inf_{\alpha}\|R_\alpha(Tf^\dagger+\xi)-f^\dagger\|\leq\inf_{\alpha}\sup_{\|\xi\|\leq\delta}\|R_\alpha(Tf^\dagger+\xi)-f^\dagger\|, \tag{19}$$

the more difficult implication $(ii)\Rightarrow(i)$ in Theorem 4.1 is weaker than in [1]. However, we do not have to impose additional assumptions relating the regularization method and the index function as required in [1]. Let us now state conditions under which a reverse inequality to (19) holds true:

Lemma 4.2. *Under Assumption 3.2 the estimate*

$$\inf_{0<\alpha\leq\overline{\alpha}}\sup_{\|\xi\|\leq\delta}\left\|R_\alpha(Tf^\dagger+\xi)-f^\dagger\right\|\leq 2\sqrt{2}\sup_{\|\xi\|\leq\delta}\inf_{0<\alpha\leq\overline{\alpha}}\left\|R_\alpha(Tf^\dagger+\xi)-f^\dagger\right\| \tag{20}$$

holds true for all

$$\delta\in\Delta(f^\dagger):=\left\{\frac{\|r_\alpha(T^*T)f^\dagger\|}{\|R_\alpha\|}:0<\alpha<\overline{\alpha}\right\} \tag{21}$$

This set has the following properties:

(i) *If $q_\alpha(\lambda)$ is continuous in α with $\overline{\alpha}=\infty$ for all $\lambda\in\sigma(T^*T)$ and $f^\dagger\neq 0$, then $\Delta(f^\dagger)=(0,\infty)$.*
(ii) *If $E_\alpha f^\dagger\neq 0$ for all $\alpha>0$, then 0 is always a cluster point of $\Delta(f^\dagger)$.*
(iii) *For Landweber iteration with $\mu\|T^*T\|<1$ and Lardy's method and $f^\dagger\neq 0$ the size of the gaps of $\Delta(f^\dagger)$ on a logarithmic scale is bounded by $\ln\gamma$ with*

$$\gamma:=\sup\left\{\frac{b}{a}:a,b\in\overline{\Delta(f^\dagger)}\wedge 0<a<b\wedge(a,b)\cap\Delta(f^\dagger)=\emptyset\right\}<\infty \tag{22}$$

Proof. For $\delta\in\Delta(f^\dagger)$ there exists $\alpha'=\alpha'(\delta,f^\dagger)$ such that

$$\left\|r_{\alpha'}(T^*T)f^\dagger\right\|=\|R_{\alpha'}\|\,\delta.$$

By the definition of the operator norm, for each $\epsilon>0$ there exists a noise vector ξ' with $\|\xi'\|\leq\delta$ such that $\|R_{\alpha'}\xi'\|\geq(1-\epsilon)\|R_{\alpha'}\|\delta$ (if T is compact we may choose $\epsilon=0$). We claim that ξ' (depending on α' and $f^\dagger$) can be chosen such that

$$\left\langle r_\alpha(T^*T)f^\dagger,R_\alpha\xi'\right\rangle\geq 0\qquad\text{for all }\alpha\in(0,\overline{\alpha}]. \tag{23}$$

Let $T=U(T^*T)^{1/2}$ be the polar decomposition of T with a unitary operator $U:\mathbb{X}\to\overline{R(T)}\subset\mathbb{Y}$ (recall that T is assumed to be injective). As $\overline{R(T)}\supset N(R_\alpha)^\perp$, we may assume w.l.o.g. that $\xi'\in\overline{R(T)}$. Hence, (23) is equivalent to $\left\langle r_\alpha(T^*T)f^\dagger,q_\alpha(T^*T)(T^*T)^{1/2}\xi''\right\rangle\geq 0$ with $\xi'=U\xi''$. By the Halmos version of the spectral theorem [18], T^*T is unitarily equivalent to a multiplication operator $M_\lambda:L^2(\Omega,\mu)\to L^2(\Omega,\mu)$, $(M_\lambda h)(z)=\lambda(z)h(z)$, $z\in\Omega$ on a locally compact space Ω with positive Borel measure μ and a non-negative function $\lambda \in L^\infty(\Omega, \mu)$, i.e. $T^*T = W^* M_\lambda W$ for some unitary operator $W : \mathbb{X} \to L^2(\Omega, \mu)$ (if T is compact, μ may be chosen as counting measure on $\Omega = \mathbb{N}$). It follows that

$$\left\langle r_\alpha(T^*T)f^\dagger, R_\alpha \xi' \right\rangle = \int_\Omega r_\alpha\left(\lambda(z)\right) \left(W f^\dagger\right)(z)\, q_\alpha\left(\lambda(z)\right) \sqrt{\lambda(z)}\, \left(W\xi''\right)(z)\, d\mu(z).$$

By Assumption 3.2 we have $r_\alpha \geq 0$ and $q_\alpha \geq 0$ for all $\alpha > 0$ (see (14)). Therefore, the right hand side of the last equation is non-negative if $(Wf^\dagger)(z)(W\xi'')(z) \geq 0$ for μ-almost all $z \in \Omega$. This may be achieved by replacing $(W\xi'')(z)$ by $s(z)(W\xi'')(z)$ with a measurable function $s : \Omega \to \{-1, 1\}$. This shows that (23) holds true if ξ' is replaced by $UW^*(s \cdot (W\xi''))$.

With this choice of α' and ξ' we can estimate

$$\begin{aligned}
\inf_{\alpha>0} \sup_{\|\xi\|\leq\delta} \left\| R_\alpha(Tf^\dagger + \xi) - f^\dagger \right\| &\leq \inf_{\alpha>0} \sup_{\|\xi\|\leq\delta} \left[\left\| r_\alpha(T^*T)f^\dagger \right\| + \|R_\alpha \xi\| \right] \\
&\leq \left\| r_{\alpha'}(T^*T)f^\dagger \right\| + \sup_{\|\xi\|\leq\delta} \|R_{\alpha'}\xi\| \\
&\leq \left\| r_{\alpha'}(T^*T)f^\dagger \right\| + \frac{1}{1-\epsilon} \|R_{\alpha'}\xi'\| \\
&\leq \frac{2}{1-\epsilon} \inf_{0<\alpha\leq\overline{\alpha}} \left[\left\| r_\alpha(T^*T)f^\dagger \right\| + \|R_\alpha \xi'\| \right]
\end{aligned}$$

since the first term in the last line is monotonically increasing and the second term is monotonically decreasing in α. As $(\|x\| + \|y\|)^2 \leq 2\|x+y\|^2$ for $\langle x, y\rangle \geq 0$, we obtain via (23) that

$$\begin{aligned}
\inf_{\alpha>0} \sup_{\|\xi\|\leq\delta} \left\| R_\alpha(Tf^\dagger + \xi) - f^\dagger \right\| &\leq \frac{2\sqrt{2}}{1-\epsilon} \inf_{0<\alpha\leq\overline{\alpha}} \left\| R_\alpha(Tf^\dagger + \xi') - f^\dagger \right\| \\
&\leq \frac{2\sqrt{2}}{1-\epsilon} \sup_{\|\xi\|\leq\delta} \inf_{0<\alpha\leq\overline{\alpha}} \left\| R_\alpha(Tf^\dagger + \xi) - f^\dagger \right\|.
\end{aligned}$$

As $\epsilon > 0$ was arbitrary, we have proven (20). Let us now show the properties of $\Delta(f^\dagger)$:

(i) If $q_\alpha(\lambda)$ and $r_\alpha(\lambda)$ are continuous in α, then $\|r_\alpha(T^*T)f^\dagger\|$ is continuous in α by Lebesgue's Dominated Convergence Theorem, and so is $\alpha \mapsto \|R_\alpha\|$. Moreover, $\lim_{\alpha\to 0} \|r_\alpha(T^*T)f^\dagger\| = 0$ as T is assumed to be injective, $\alpha \mapsto \|R_\alpha\|$ is decreasing, by (17) we have $\|R_\alpha\| \to 0$ as $\alpha \to \overline{\alpha}$, and by (14) we have $\|r_\alpha(T^*T)f^\dagger\| \leq \|f^\dagger\|$ for all α. As $f^\dagger \neq 0$, the case $r_\alpha(T^*T)f^\dagger = 0$ for all $\alpha > 0$ can be excluded by noting that

$$r_\alpha(\lambda) > 0 \qquad \text{for } \alpha > C_1\lambda \tag{24}$$

since $q_\alpha(\lambda) \leq C_1/\alpha < 1/\lambda$. This shows that $\|r_\alpha(T^*T)f^\dagger\|/\|R_\alpha\|$ tends to 0 as $\alpha \to 0$ and to ∞ as $\alpha \to \overline{\alpha}$. Now $\Delta(f^\dagger) = (0, \infty)$ follows from continuity and the intermediate value theorem.

(ii) If $q_\alpha(\lambda)$ is not continuous or $\overline{\alpha} \neq \infty$, we still have the same limiting behaviour of $\|r_\alpha(T^*T)f^\dagger\|/\|R_\alpha\|$ for $\alpha \to 0$. However, we have to exclude the case $r_\alpha(T^*T)f^\dagger =$

0 for all α in some neighborhood of 0 to ensure that 0 is a cluster point of $\Delta(f^\dagger)$. This is achieved by (24) and the assumption $E_\alpha f^\dagger \neq 0$ for all $\alpha > 0$.

(iii) For Landweber iteration and Lardy's method we have

$$\Delta(f^\dagger) = \left\{\delta_n(f^\dagger) := \frac{\|r_{1/n}(T^*T)f^\dagger\|}{\|R_{1/n}\|} : n \in \mathbb{N}\right\} \qquad \text{and} \qquad \gamma = \sup_{n\in\mathbb{N}} \frac{\delta_n(f^\dagger)}{\delta_{n+1}(f^\dagger)}.$$

Using (17) we can bound quotients of the denominators of $\delta_n(f^\dagger)$ by

$$\begin{aligned}\|R_{1/(n+1)}\| &= \sup_{\lambda\in\sigma(T^*T)} \frac{1 - r_{1/(n+1)}(\lambda)}{\sqrt{\lambda}} \leq \sup_{\lambda\in\sigma(T^*T)} \frac{1 - r_{1/(n+1)}(\lambda)}{1 - r_{1/n}(\lambda)} \sup_{\lambda\in\sigma(T^*T)} \frac{1 - r_{1/n}(\lambda)}{\sqrt{\lambda}} \\ &\leq \sup_{\lambda\in[0,\|T^*T\|]} \frac{1 - r_{1/(n+1)}(\lambda)}{1 - r_{1/n}(\lambda)} \|R_{1/n}\|.\end{aligned}$$

Now setting $x = (1 - \mu\lambda)$ and $x = \beta/(\beta + \lambda)$ for Landweber iteration and Lardy's method respectively we obtain the bound

$$\frac{\|R_{1/(n+1)}\|}{\|R_{1/n}\|} \leq \sup_{x\in[0,1]} \frac{1 - x^{n+1}}{1 - x^n} = \sup_{x\in[0,1]} \left(x + \frac{1 - x}{1 - x^n}\right) \leq 2.$$

For Landweber iteration quotients of enumerators of $\delta_n(f^\dagger)$ are bounded by

$$\frac{\|r_{1/n}(T^*T)f^\dagger\|^2}{\|r_{1/(n+1)}(T^*T)f^\dagger\|^2} = \frac{\int_0^{\|T^*T\|} (1 - \mu\lambda)^{2n} \,\mathrm{d}\|E_\lambda f^\dagger\|^2}{\int_0^{\|T^*T\|} (1 - \mu\lambda)^{2n+2} \,\mathrm{d}\|E_\lambda f^\dagger\|^2} \leq \frac{1}{(1 - \mu \|T^*T\|)^2}$$

since $(1 - \mu\lambda) \geq 1 - \mu \|T^*T\|$ for all $\lambda \leq \|T^*T\|$. Similar for Lardy's method we use $\beta/(\beta + \lambda) \geq \beta/(\beta + \|T^*T\|)$ for all $\lambda \leq \|T^*T\|$ to obtain

$$\frac{\|r_{1/n}(T^*T)f^\dagger\|^2}{\|r_{1/(n+1)}(T^*T)f^\dagger\|^2} = \frac{\int_0^{\|T^*T\|} \left(\frac{\beta}{\beta+\lambda}\right)^{2n} \mathrm{d}\|E_\lambda f^\dagger\|^2}{\int_0^{\|T^*T\|} \left(\frac{\beta}{\beta+\lambda}\right)^{2n+2} \mathrm{d}\|E_\lambda f^\dagger\|^2} \leq \left(1 + \frac{\|T^*T\|}{\beta}\right)^2.$$

This shows that γ is finite in both cases. □

The difference between our Theorem 4.1 and the corresponding results in [1,17] is analogous to the difference between the concepts of weakly and strongly quasioptimal parameter choice rules as introduced by Raus & Hämarik [20]. They called a parameter choice rule $\alpha_* : [0, \infty) \times \mathbb{Y} \to [0, \infty)$ *weakly quasioptimal* (or simply quasioptimal) for the regularization method $\{R_\alpha\}$ if there exists a constant $C > 0$ such that

$$\|R_{\alpha_*(\delta,g^{\mathrm{obs}})}g^{\mathrm{obs}} - f^\dagger\| \leq C \inf_{\alpha>0} \sup_{\|\xi\|\leq\delta} \|R_\alpha(Tf^\dagger + \xi) - f^\dagger\| + \mathcal{O}(\delta) \tag{25}$$

for all $f^\dagger \in \mathbb{X}$ and all $g^{\mathrm{obs}} \in \mathbb{Y}$ with $\|g^{\mathrm{obs}} - Tf^\dagger\| \leq \delta$ as $\delta \to 0$. (Our formulation of this definition slightly differs from that in [20], but it is equivalent due to the arguments at the beginning of part 2 of the proof of Theorem 4.1.) A parameter choice rule α_* is called *strongly quasioptimal* if there exists $C > 0$ such that

$$\|R_{\alpha_*(\delta,g^{\mathrm{obs}})}g^{\mathrm{obs}} - f^\dagger\| \leq C \sup_{\|\xi\|\leq\delta} \inf_{\alpha>0} \|R_\alpha(Tf^\dagger + \xi) - f^\dagger\| + \mathcal{O}(\delta) \tag{26}$$

for all $f^\dagger \in \mathbb{X}$ and all $g^{\text{obs}} \in \mathbb{Y}$ with $\|g^{\text{obs}} - Tf^\dagger\| \le \delta$. Lemma 4.2 shows that the notions of weak and strong quasioptimality coincide for continuous regularization methods. In [20] it is shown that the discrepancy principle is strongly quasioptimal for regularization methods of infinite qualification such as Landweber iteration, Lardy's method or spectral cut-off (but it is not even weakly quasioptimal for Tikhonov regularization and iterated Tikhonov regularization). For iterated Tikhonov regularization the Raus-Gfrerer rule is weakly quasioptimal by results in [19], and hence by Lemma 4.2 also strongly quasioptimal. Moreover, Lepskiĭ's rule was shown to be weakly quasioptimal for all considered methods ([20]), and hence it is strongly quasioptimal for (iterated) Tikhonov regularization, Showalter's method, and modified spectral cut-off by Lemma 4.2. The Monotone Error Rule is strongly quasioptimal for Landweber iteration and Lardy's method ([20]). In most cases the constant C in (25) or (26) can be given explicitly.

Theorem 4.3. *Suppose Assumption 3.2 holds true, let α_* be weakly quasioptimal parameter choice rule, let ψ_κ be concave, and assume that (16) holds true. Then the following statements are equivalent for any $f^\dagger \in \mathbb{X}$ for which $\Delta(f^\dagger)$ satisfies (22):*

(i) $\sup_{\overline{\alpha}\ge\alpha>0} \frac{1}{\kappa(\alpha)^2}\|r_\alpha(T^*T)f^\dagger\|^2 < \infty$.

(ii) *For any finite $\delta_0 > 0$ we have*

$$\sup_{\delta\in(0,\delta_0]} \frac{1}{\psi_\kappa(\delta^2)} \sup_{\|\xi\|\le\delta} \|R_{\alpha_*(\delta,Tf^\dagger+\xi)}(Tf^\dagger+\xi) - f^\dagger\|^2 < \infty.$$

Proof. *(i) ⇒ (ii):* Using Theorem 4.1 and the definition of weak quasioptimality (25) we see that there exists a constant C such that for all $\delta > 0$ the estimate

$$\sup_{\|\xi\|\le\delta} \|R_{\alpha_*(\delta,Tf^\dagger+\xi)}(Tf^\dagger+\xi) - f^\dagger\|^2 \le C\left(\psi_\kappa(\delta^2)+\delta^2\right)$$

holds true. Since ψ_κ is concave $\lim_{t\to0} t/\psi_\kappa(t)$ is bounded and hence we obtain (ii) for any finite $\delta_0 > 0$.

(ii) ⇒ (i): By Lemma 4.2 we have

$$\begin{aligned}\inf_{\overline{\alpha}\ge\alpha>0} \sup_{\|g^{\text{obs}}-Tf^\dagger\|\le\delta} \|R_\alpha(g^{\text{obs}}) - f^\dagger\|^2 &\le 8 \sup_{\|g^{\text{obs}}-Tf^\dagger\|\le\delta} \inf_{\overline{\alpha}\ge\alpha>0} \|R_\alpha(g^{\text{obs}}) - f^\dagger\|^2 \\ &\le 8 \sup_{\|g^{\text{obs}}-Tf^\dagger\|\le\delta} \|R_{\alpha_*(\delta,g^{\text{obs}})}(g^{\text{obs}}) - f^\dagger\|^2 \le 8C(f^\dagger,\delta_0)\psi_\kappa(\delta)\end{aligned}$$

for all $\delta \in \Delta(f^\dagger)\cap[0,\delta_0]$. Now choose $\delta_0 = \Theta_\kappa(2B\beta/(1-C_3)^2)$ with $\beta = \overline{\alpha}\wedge\|T^*T\|$ and first assume that $\Delta(f^\dagger)\cap(0,\delta_0] = (0,\delta_0]$. Then we obtain $\sup_{\alpha\in(0,\beta]} \frac{1}{\kappa(\alpha)^2}\|r_\alpha(T^*T)f^\dagger\|^2 < \infty$ following the proof of Theorem 4.1. As $\|r_\alpha(T^*T)f^\dagger\|^2 \le \|f^\dagger\|^2$ for all $\alpha > 0$, this is equivalent to (i).

If $\Delta(f^\dagger)$ has gaps, but satisfies (22), then for each $\delta \in (0,\delta_0]$ we can find $\underline{\delta} \in [\delta/\gamma,\delta]$ with $\underline{\delta} \in \Delta(f^\dagger)$. By the concavity of ψ_κ we have $\psi_\kappa(\delta)/\psi_\kappa(\underline{\delta}) \le \gamma$. Therefore, replacing the supremum over $\delta \in (0,\delta_0]\cap\Delta(f^\dagger)$ by the supremum over $\delta \in (0,\delta_0]$ increases the value at most by a factor γ. □

5. Converse results for white noise

We now want to prove a theorem similar to Theorem 4.1 for the white noise error model (2) using the expected square error as error measure. By the bias-variance decomposition this equals

$$\mathbf{E}\left[\left\|\widehat{f}_\alpha - f^\dagger\right\|_{\mathbb{X}}^2\right] = \left\|\mathbf{E}\left[\widehat{f}_\alpha\right] - f^\dagger\right\|_{\mathbb{X}}^2 + \varepsilon^2\mathbf{E}\left[\|R_\alpha W\|^2\right]. \tag{27}$$

By Theorem 3.3 the bias can be controlled by assuming that $f^\dagger \in \mathbb{X}_\kappa^T$. The variance is given by $\varepsilon^2\mathbf{E}\left[\|R_\alpha W\|\right]^2 = \varepsilon^2\mathrm{trace}(R_\alpha^* R_\alpha)$, i.e. as opposed to the deterministic the effect of the noise is not described by the maximum, but by the sum of the eigenvalues of $R_\alpha^* R_\alpha$. Often the sum grows faster than the maximum as $\alpha \to 0$, and the specific rate depends not only on the regularization methods, but also on the eigenvalue distribution of the operator. We will assume that there exists a constant $D \geq 1$ and a monotonically decreasing function $v \in C((0,\infty))$ such that

$$\frac{1}{D}v(\alpha)^2 \leq \mathbf{E}\left[\|R_\alpha W\|^2\right] \leq Dv(\alpha)^2 \qquad \forall \overline{\alpha} \geq \alpha > 0 \tag{28a}$$

with limits $\lim_{t\to 0} v(t) = \infty$ and $\lim_{t\to\infty} v(t) = 0$. Note the in the deterministic case, i.e. for $\mathbf{E}\left[\|R_\alpha W\|^2\right]$ replaced by $\|R_\alpha\|^2$, we could simply choose $v(\alpha) = c/\sqrt{\alpha}$. Moreover, we will assume that $v(\alpha)$ does not grow faster than polynomially as $\alpha \to 0$, or equivalently, that the inverse function $v^{-1} : (0,\infty) \to (0,\infty)$ does not decay faster than polynomially at infinity in the sense that there exists $p \geq 1$ such that

$$v^{-1}(rt) \geq r^{-q} v^{-1}(t) \tag{28b}$$

for all $t > 0$ and $r \geq 1$.

It was shown in [5] that $\mathbf{E}\left[\|R_\alpha W\|^2\right] \sim \mathbf{E}\left[\|(T(I - E_\alpha^{T^*T}))^\dagger\|^2\right]$ under certain conditions, and explicit expressions for v have been derived.

Theorem 5.1. *Let Assumption 3.2 and (28) hold true and define*

$$\psi_{\kappa,v}(t) := \kappa\left(\Theta_{\kappa,v}^{-1}\left(\sqrt{t}\right)\right)^2 \qquad \textit{with} \qquad \Theta_{\kappa,v}(\alpha) := \frac{\kappa(\alpha)}{v(\alpha)}.$$

Moreover, assume that κ satisfies (16). Then for $f^\dagger \in \mathbb{X}$ the following statements are equivalent:

(i) $A := \sup_{0<\alpha\leq\overline{\alpha}} \frac{1}{\kappa(\alpha)^2}\|r_\alpha(T^*T)f^\dagger\|_{\mathbb{X}}^2 < \infty$
(ii) $B := \sup_{0<\varepsilon\leq\Theta_{\kappa,v}(\overline{\alpha})} \frac{1}{\psi_{\kappa,v}(\varepsilon^2)} \inf_{0<\alpha\leq\overline{\alpha}} \mathbf{E}[\|R_\alpha(Tf^\dagger + \varepsilon W) - f^\dagger\|_{\mathbb{X}}^2] < \infty.$

More precisely,

$$B \leq A + D \qquad \textit{and} \qquad A \leq B \vee B(2BD)^{pq} \vee \frac{\|f^\dagger\|^2}{\kappa\left(v^{-1}(\sqrt{2BD}v(\overline{\alpha}))\right)}.$$

Proof. (i) ⇒ (ii): Set $\widehat{f}_\alpha := R_\alpha(Tf^\dagger + \varepsilon W)$. By (i) and Theorem 3.3 we can bound $\|\mathbf{E}[\widehat{f}_\alpha] - f^\dagger\|_{\mathbb{X}}^2 = \|r_\alpha(T^*T)f^\dagger\|_{\mathbb{X}}^2 \leq A\kappa(\alpha)^2$, and by assumption (28a) we can estimate $\mathbf{E}[\|R_\alpha W\|^2] \leq Dv^2(\alpha)$. Hence,

$$\mathbf{E}\left[\left\|\widehat{f}_\alpha - f^\dagger\right\|_{\mathbb{X}}^2\right] \leq A\kappa^2(\alpha) + D\varepsilon^2 v^2(\alpha).$$

The minimum over the right hand side is approximately attained if $\kappa(\alpha) = \varepsilon v(\alpha)$ or equivalently if $\alpha = \Theta_{\kappa,v}^{-1}(\varepsilon)$. The equality $\kappa(\alpha) = \varepsilon v(\alpha)$ implies in particular that

$$\psi_{\kappa,v}(\varepsilon^2) = \varepsilon^2 v(\Theta_{\kappa,v}^{-1}(\varepsilon))^2. \tag{29}$$

Therefore, for all $\varepsilon > 0$ we obtain

$$\inf_{0<\alpha\leq\overline{\alpha}} \mathbf{E}\left[\left\|\widehat{f}_\alpha - f^\dagger\right\|_{\mathbb{X}}^2\right] \leq [A+D]\,\kappa\left(\Theta_{\kappa,v}^{-1}(\varepsilon)\right)^2$$

and can choose $B = A + D$.

(ii) ⇒ (i): Using again (27) and the lower bound on the variance in (28a) we obtain

$$\mathbf{E}\left[\left\|\widehat{f}_\alpha - f^\dagger\right\|_{\mathbb{X}}^2\right] \geq \left\|\mathbf{E}\left[\widehat{f}_\alpha\right] - f^\dagger\right\|_{\mathbb{X}}^2 + \frac{1}{D}v(\alpha) = \left\|r_\alpha(T^*T)f^\dagger\right\|_{\mathbb{X}}^2 + \frac{\varepsilon^2}{D}v(\alpha)^2.$$

Because the first term is increasing and the second term is decreasing in α, we obtain

$$B\psi_{\kappa,v}(\varepsilon^2) \geq \inf_{\alpha>0}\left[\left\|r_\alpha(T^*T)f^\dagger\right\|_{\mathbb{X}}^2 + \frac{\varepsilon^2}{D}v(\alpha)^2\right] \geq \left\|r_{\alpha_*}(T^*T)f^\dagger\right\|_{\mathbb{X}}^2 \wedge \frac{\varepsilon^2}{D}v(\alpha_*)^2. \tag{30}$$

for any $\alpha_* \in (0,\overline{\alpha}]$. We will choose

$$\alpha_*(\varepsilon) = v^{-1}\left(\sqrt{2BD}\,v\left(\Theta_{\kappa,v}^{-1}(\varepsilon)\right)\right). \tag{31}$$

Using (29) we see that the second term in the minimum in (30) equals twice the left hand side:

$$\frac{\varepsilon^2}{D}v(\alpha_*)^2 = 2B\varepsilon^2 v(\Theta_{\kappa,v}^{-1}(\varepsilon))^2 = 2B\psi_{\kappa,v}(\varepsilon^2),$$

Therefore, $\left\|r_{\alpha_*}(T^*T)f^\dagger\right\|_{\mathbb{X}}^2 \geq \frac{\varepsilon^2}{D}v(\alpha_*)^2$ leads to a contradiction, i.e. the minimum in (30) is attained at the first argument. We obtain

$$\left\|r_{\alpha_*}(T^*T)f^\dagger\right\|_{\mathbb{X}}^2 \leq B\psi_{\kappa,v}(\varepsilon^2) = B\psi_{\kappa,v}\left(\left(\Theta_{\kappa,v}\left(v^{-1}\left(\frac{v(\alpha_*)}{\sqrt{BD}}\right)\right)\right)^2\right),$$

where we have solved (31) for ε in the second step. Abbreviating $z := v^{-1}\left(v(\alpha_*)/\sqrt{2BD}\right)$ and using (29) again, we find

$$\begin{aligned}\left\|r_{\alpha_*}(T^*T)f^\dagger\right\|_{\mathbb{X}}^2 &\leq B\psi_{\kappa,v}\left((\Theta_{\kappa,v}(z))\right)^2 = B\left(\Theta_{\kappa,v}(z)\right)^2 v\left(\Theta_{\kappa,v}^{-1}\left(\Theta_{\kappa,v}(z)\right)\right)^2 \\ &= B\kappa(z)^2 \leq B\kappa\left(\left(1 \vee (2BD)^{q/2}\right)\alpha_*\right)^2 \leq \left(B \vee B(2BD)^{pq}\right)\kappa(\alpha_*)^2\end{aligned}$$

for all $\alpha \in (0,\overline{\alpha}]$ defined by (31) using (28b).

For $\alpha \in (0, \overline{\alpha}]$ not of the given form note that $\alpha \geq v^{-1}(\sqrt{2BD}v(\overline{\alpha}))$ and using mononicity we obtain for these α

$$\frac{1}{\kappa(\alpha)^2}\left\|r_\alpha(T^*T)f^\dagger\right\|_{\mathbb{X}}^2 \leq \frac{\|f^\dagger\|^2}{\kappa\left(v^{-1}(\sqrt{2BD}v(\overline{\alpha}))\right)}$$

showing boundedness for all $\alpha \in (0, \overline{\alpha}]$. □

Remark 5.2. If assumption (28a) is relaxed to

$$v_-(\alpha)^2 \leq \mathbf{E}\left[\|R_\alpha W\|^2\right] \leq v_+(\alpha)^2 \qquad \forall \overline{\alpha} \geq \alpha > 0 \tag{32}$$

where possibly $\lim_{\alpha\to 0}(v_+/v_-)(\alpha) = \infty$, and v_- satisfies (28b), then it can be seen by inspection of the proof that

$$\text{Theorem 5.1(i)} \;\Rightarrow\; \sup_{0<\varepsilon\leq\Theta_{\kappa,v_+}(\overline{\alpha})} \frac{1}{\psi_{\kappa,v_+}(\varepsilon^2)} \inf_{0<\alpha\leq\overline{\alpha}} \mathbf{E}\left[\|R_\alpha(Tf^\dagger + \varepsilon W) - f^\dagger\|_{\mathbb{X}}^2\right] < \infty,$$

$$\text{Theorem 5.1(i)} \;\Leftarrow\; \sup_{0<\varepsilon\leq\Theta_{\kappa,v_-}(\overline{\alpha})} \frac{1}{\psi_{\kappa,v_-}(\varepsilon^2)} \inf_{0<\alpha\leq\overline{\alpha}} \mathbf{E}\left[\|R_\alpha(Tf^\dagger + \varepsilon W) - f^\dagger\|_{\mathbb{X}}^2\right] < \infty.$$

This is relevant for operators T with exponentially decaying singular values. Whereas for polynomial decay assumption (28a) can be verified using results from [5], for singular values with asympototic behaviour $\sigma_j(T) \sim \exp(-cj^\beta)$ with $c, \beta > 0$ one can only (easily) verify the relaxed condition (32) with

$$v_-(\alpha) = c_-\alpha^{-1/2} \qquad \text{and} \qquad v_+(\alpha) = c_+\alpha^{-1/2-\tau}$$

for any $\tau > 0$ and some $c_-, c_+ > 0$. However, for such operators (i) is typically satisfied only for logarithmic functions $\kappa(\alpha) = (-\ln\alpha)^{-p}$ with some $p > 0$ for $f^\dagger$ of finite smoothness. In this case one has

$$\psi_{\kappa,v_+}(t) = (-\ln t)^{-2p}(1 + o(1)), \qquad t \to 0$$

independent of the choice of $\tau \in [0, \infty)$ (see [15]). Therefore, the equivalence in Theorem 5.1 still holds true with either $v = v_-$ or $v = v_+$.

6. Besov spaces as maxisets

We have seen in the previous sections that convergence rates of ψ_κ to a true solution $f^\dagger$ for regularization methods are completely characterized by $f^\dagger \in \mathbb{X}_\kappa^T$. Andreev [3] showed that these spaces coincide with K-interpolation spaces with equivalent norms. Recall that for a Banach space $\mathbb{Z} \subset \mathbb{X}$, which is continuously embedded in $\mathbb{X}$ the K-functional is defined by

$$K_t(f) := \inf_{g\in\mathbb{Z}} \left(\|f-g\|_{\mathbb{X}}^2 + t^2\|g\|_{\mathbb{Z}}^2\right)^{1/2}.$$

For $\nu \in (0,1)$ the K-interpolation space with fine index ∞ is defined by

$$(\mathbb{X}, \mathbb{Z})_{\nu,\infty} := \{f \in \mathbb{X} : \|f : (\mathbb{X}, \mathbb{Z})_{\nu,\infty}\| < \infty\} \quad \text{where} \quad \|f : (\mathbb{X}, \mathbb{Z})_{\nu,\infty}\| := \sup_{t>0} t^{-\nu} K_t(f).$$

Here we temporarily switch to a different norm notation because of the numereous indices. It can be shown that $(\mathbb{X},\mathbb{Z})_{\nu,\infty}$ with this norm is a Banach space. If $\mathbb{Z} = (S^*S)^k(\mathbb{X})$ for a bounded linear operator $S : \mathbb{X} \to \mathbb{Y}$ and some $k \in \mathbb{N}$ with norm $\|f\|_{\mathbb{Z}} := \|(S^*S)^{-k}f\|_{\mathbb{X}}$, Andreev [3] showed that

$$\mathbb{X}^S_{\mathrm{id}^{k\nu}} = (\mathbb{X},(S^*S)^k(\mathbb{X}))_{\nu,\infty} \tag{33}$$

for $\nu \in (0,1)$ with

$$\sqrt{1-\nu}\left\|f : (\mathbb{X},(S^*S)^k(\mathbb{X})_{\nu,\infty}\right\| \le \left\|f : \mathbb{X}^S_{\mathrm{id}^{k\nu}}\right\| \le (1-\nu)^{1-\nu}\,\nu^\nu \left\|f : (\mathbb{X},(S^*S)^k(\mathbb{X})_{\nu,\infty}\right\|.$$

We further recall that the K-interpolation of certain Sobolev spaces yields Besov spaces. In particular,

$$(L^2(\mathcal{M}), H^k(\mathcal{M}))_{\nu,\infty} = B^{k\nu}_{2,\infty}(\mathcal{M}) \tag{34}$$

if $\mathcal{M}$ is a smooth Riemannian manifold with Laplace-Beltrami operator Δ satisfying Assumption 6.1 below and $H^k(\mathcal{M}) := (I-\Delta)^{-k/2}(L^2(\mathcal{M}))$ with norm $\|f\|_{H^k} := \|(I-\Delta)^{k/2}f\|_{L^2}$ (see [22, Chapter 7]).

Assumption 6.1. Let $\mathcal{M}$ be a connected smooth Riemanian manifold. Let $\mathcal{M}$

- be complete,
- have an injectivity radius $r > 0$ and
- a bounded geometry.

Here completeness means that all geodesics are infinitely extendable, the injectivity radius refers to the size of the domains in which the exponential map is bijective, and bounded geometry means that the determinant of the Riemannian metric is bounded from below by a positive constant and all its derivatives are bounded from above (see [22] for further discussions). Important examples of such manifolds include $\mathbb{R}^n$ and compact manifolds without boundaries.

In the following we will consider operators $T\colon \mathbb{X} = L^2(\mathcal{M}) \to \mathbb{Y}$ such that

$$T^*T = \Lambda(-\Delta), \tag{35}$$

where Λ fulfills the following conditions:

Assumption 6.2. Let $\Lambda : [0,\infty) \to (0,\infty)$ such that

- Λ is continuous,
- $\Lambda|_{[t_0,\infty)}$ is strictly decreasing for some $t_0 \ge 0$,
- $\Lambda(\mu) \to 0$ for $\mu \to \infty$.

Our aim of this section is to prove the following theorem:

Theorem 6.3. *Let $\mathcal{M}$ fulfill Assumption 6.1, Λ fulfill Assumption 6.2 and $s > 0$. Let $T : L^2(\mathcal{M}) \to \mathbb{Y}$ be of the form (35) and define*

$$\kappa(\alpha) := \begin{cases} 0, & \text{if } \alpha = 0\\ \left(\Lambda|_{[t_0,\infty)}{}^{-1}(\alpha)\right)^{-1/2}, & \text{if } \alpha \in ((0,\Lambda(t_0)]),\\ t_0^{-1/2}, & \text{if } \alpha > \Lambda(t_0). \end{cases}$$

Then $\mathbb{X}^T_{\kappa^s} = B^s_{2,\infty}(\mathcal{M})$ *with equivalent norms.*

Proof. We introduce the operator $S := \kappa(T^*T)^{1/2} : L^2(\mathcal{M}) \to L^2(\mathcal{M}))$ such that

$$S^*S = \kappa(T^*T) = (\kappa \circ \Lambda)(-\Delta).$$

As $(\kappa \circ \Lambda)(t) = t^{-1/2}$ for $t \geq t_0$ and $\inf_{0\leq t\leq t_0}(\kappa \circ \Lambda)(t) > 0$ by continuity, we have

$$(S^*S)^k(L^2(\mathcal{M})) = H^k(\mathcal{M}) \tag{36}$$

with equivalent norms for all $k \in \mathbb{N}$. Using the substitution $t = \kappa(\alpha)$ we obtain

$$\left\| f : \mathbb{X}^S_{\mathrm{id}^s} \right\| = \sup_{0<t\leq 1/t_0} t^{-s} \left\| E^{S^*S}_t f \right\| = \sup_{\alpha\in(0,\Lambda(t_0)]} \frac{1}{\kappa(\alpha)^s} \left\| E^{S^*S}_{\kappa(\alpha)} f \right\| = \sup_{\alpha\in(0,\Lambda(t_0)]} \frac{1}{\kappa(\alpha)^s} \left\| E^{T^*T}_\alpha f \right\|.$$

As $E^{T^*T}_\alpha(f) = E^{T^*T}_{\Lambda(0)}(f) = f$ for $\alpha > \Lambda(0)$, this shows that the norms $\left\| f : \mathbb{X}^S_{\mathrm{id}^s} \right\|$ and $\left\| f : \mathbb{X}^T_{\kappa^s} \right\|$ are equivalent. Choosing $k \in \mathbb{N}$ with $k > s$ and using (33), (36) and (34) we obtain

$$\mathbb{X}^T_{\kappa^s} = \mathbb{X}^S_{\mathrm{id}^s} = (L^2(\mathcal{M}), (S^*S)^k(\mathcal{M}))_{s/k,\infty} = B^s_{2,\infty}(\mathcal{M})$$

with equivalent norms. □

7. Examples

In this section we want to apply our results to some examples. The examples are taken from [10] and complement the results there.

7.1. **Operators in Sobolev scales.** In the following we describe a fairly general class of problems. It contains convolution operators (if $\mathcal{M} = \mathbb{R}^d$ or $\mathcal{M} = (\mathbb{S}^1)^d$), for which the convolution kernel has a certain type of singularity at 0, boundary integral operators, injective elliptic pseudo-differential operators, and compositions of such operators.

Theorem 7.1. *Let* $\mathcal{M}$ *be a d-dimensional manifold satisfying Assumption 6.1, and let* T *be an operator which is a times smoothing* ($a > d/2$) *in the sense that* $T : H^s(\mathcal{M}) \to H^{s+a}(\mathcal{M})$ *is well-defined, bounded and has a bounded inverse for all* $s \in \mathbb{R}$. *We will consider* T *as an operator from* $L^2(\mathcal{M})$ *into itself, i.e.* $\mathbb{X} = \mathbb{Y} = L^2(\mathcal{M})$. *We consider a spectral regularization method with classical qualification* $\mu_0 \geq 1$ *satisfying Assumption 3.2. Then the following statements are equivalent for all* $f^\dagger \in \mathbb{X} \setminus \{0\}$ *and* $u \in (0, a)$*:*

(i) $f^\dagger$ *satisfies a the VSC (3) with* $\psi(t) = Ct^{\frac{u}{u+a}}$ *for some* $C > 0$.

(ii) $f^\dagger \in B^u_{2,\infty}(\mathcal{M})$.

(iii) *For a quasioptimal parameter choice rule* α_* *and a regularization method for which* $\Delta(f^\dagger)$ *satisfies (22) we have*

$$\sup\{\left\|R_{\alpha_*(\delta,g^{\mathrm{obs}})}g^{\mathrm{obs}} - f^\dagger\right\|_{L^2} : \left\|g^{\mathrm{obs}} - Tf^\dagger\right\|_{L^2} \leq \delta\} = \mathcal{O}\left(\delta^{\frac{u}{u+a}}\right), \qquad \delta \to 0.$$

(iv) $$\left(\inf_{\alpha>0} \mathbf{E}\left[\left\|R_\alpha(Tf^\dagger + \varepsilon W) - f^\dagger\right\|^2_{L^2}\right]\right)^{1/2} = \mathcal{O}\left(\varepsilon^{\frac{u}{u+a+d/2}}\right), \qquad \varepsilon \to 0.$$

(ii)–(iv) are equivalent for all $u \in (0, 2a\mu_0)$*, and the assumption* $a > d/2$ *can be relaxed to* $a > 0$ *if (iv) is neglected.*

Proof. *(i)* $\Leftrightarrow f^\dagger \in \mathbb{X}^T_\kappa$*:* Note that $\psi = \psi_\kappa$ with

$$\kappa(t) = C' t^{u/2a} \qquad \text{for some } C' > 0,$$

and that the assumption $u \in (0, a)$ ensures that κ satisfies the conditions of Theorem 3.1.

$f^\dagger \in \mathbb{X}^T_\kappa \Leftrightarrow$ *(ii):* It follows from (33) and (34) that

$$\mathbb{X}^T_\kappa = (L^2(\mathcal{M}), (T^*T)(L^2(\mathcal{M})))_{u/2a,\infty} = (L^2(\mathcal{M}), H^{2a}(\mathcal{M}))_{u/2a,\infty} = B^u_{2,\infty}(\mathcal{M}).$$

$f^\dagger \in \mathbb{X}^T_\kappa \Leftrightarrow$ (37) *below*: For $u/2a < \mu_0$ Theorem 3.3 yields equivalence to

$$\sup_{\alpha>0} \alpha^{-u/2a} \|r_\alpha(T^*T)f^\dagger\| < \infty. \tag{37}$$

(37) $\Leftrightarrow$ *(iii):* This follows from Theorem 4.3.

(37) $\Leftrightarrow$ *(iv):* It has been shown in [5, §5.3] that (28a) holds true with $v(\alpha) = \alpha^{-(a+d/2)/(2a)}$. Hence, we can apply Theorem 5.1. ☐

Example 7.2. We consider a circle $\mathcal{M} = r\mathbb{S}^1 \subset \mathbb{R}^2$ with $r > 0$ and the single layer potential operator $(Tf)(x) := -\frac{1}{\pi}\int_{\mathcal{M}} \ln|x-y| f(y)\, ds(y)$. Let $f_n(r\cos t, r\sin t) := (2\pi r)^{-1/2}\exp(int)$, $n \in \mathbb{Z}$ denote the trigonometric basis of $L^2(r\mathbb{S})$. It is known (see [2, Sec. 3.3]) that $Tf_n = -1/|n| f_n$ for $n \neq 0$, and $Tf_0 = \ln(r) f_0$. Let us choose $r = \exp(1)$ for simplicity. Recall that an (equivalent) norm on $H^s(\mathcal{M})$ is given by $\|f\|^2_{H^s} = \sum_{n\in\mathbb{Z}} (1 \vee |n|)^{2s} \langle f, f_n\rangle^2$ for $s \geq 0$. W.r.t. this norm $T^\nu = (T^*T)^{u/2}$ is isometric from $H^s(\mathcal{M})$ to $H^{s+u}(\mathcal{M})$ for all $u > 0$, so the assumptions of Theorem 7.1 hold true with $a = d = 1$. Moreover, the spectral source condition $f^\dagger \in \operatorname{ran}((T^*T)^{u/2}$ is equivalent to $f^\dagger \in H^u(\mathcal{M})$ and yields the convergence rate $\mathcal{O}\left(\delta^{u/(u+1)}\right)$. The (equivalent) $\mathbb{X}^T_\kappa$-norm of $B^u_{2,\infty}(\mathcal{M})$ (with $\kappa(t) = t^u$) is given by $\|f^\dagger\|^2_{B^u_{2,\infty}} = \sup_{m\geq 0}(1 \vee m)^{2u} \sum_{|n|\geq m} \langle f^\dagger, f_n\rangle^2$. This shows that $B^u_{2,\infty}(\mathcal{M})$ is the set of $f \in L^2(\mathcal{M})$ for which the L^2-orthogonal projections onto the space of trigonometric polynomials of degree $\leq m$ converge with rate $\mathcal{O}(m^{-u})$ as $m \to \infty$. Note that

$$f^\dagger = \sum_{n\in\mathbb{Z}} (1 \vee |n|)^{-u} f_n \in B^u_{2,\infty}(\mathcal{M}) \setminus H^u(\mathcal{M})$$

for any $u > 0$, but $f^\dagger \in H^\nu(\mathcal{M})$ for $\nu < u$. Therefore, we obtain the convergence rate $\mathcal{O}(\delta^{u/(u+1)})$ for $f^\dagger$, whereas an analysis via spectral source conditions only yields rates $\mathcal{O}(\delta^{\nu/(\nu+1)})$ for $\nu \in (0, u)$. Moreover, as $\lim_{\nu\nearrow u} \|f^\dagger\|_{H^\nu} = \infty$, constants explode as $\nu \to u$.

7.2. **Backward heat equation.** Let us consider the heat equation on a manifold $\mathcal{M}$ satisfying Assumption 6.1:

$$\begin{aligned} \partial_t u &= \Delta u && \text{in } \mathcal{M} \times (0, \bar{t}) \\ u(\cdot, 0) &= f && \text{on } \mathcal{M} \end{aligned}$$

The backward heat equation is the inverse problem to estimate the initial temperature f from observations of the final temperature $g = u(\cdot, \bar{t})$. This fits into the framework (35) with the function

$$\Lambda_{\mathrm{BH}}(\mu) = \exp(-2\bar{t}\mu).$$

We obtain the following equivalence result:

Theorem 7.3. *Let $\mathcal{M}$ be a compact manifold satisfying Assumption 6.1. For spectral regularization methods satisfying Assumption (3.2) and the forward operator $T : L^2(\mathcal{M}) \to L^2(\mathcal{M})$ with $T^*T = \Lambda_{\mathrm{BH}}(-\Delta)$ of the backward heat equation the following statements for $\beta > 0$ and $f^\dagger \in L^2(\mathcal{M}) \setminus \{0\}$ are equivalent:*

(i) $f^\dagger \in B^{2\beta}_{2,\infty}(\mathcal{M})$.

(ii) *$f^\dagger$ satisfies a VSC (3) with index function $\psi(t) = C \log(3+t^{-1})^{-2\beta}$ for some $C > 0$.*

(iii) *For a quasioptimal parameter choice rule α_* and a regularization method for which $\Delta(f^\dagger)$ satisfies (22) we have*

$$\sup\{\left\|R_{\alpha_*(\delta,g^{\mathrm{obs}})}g^{\mathrm{obs}} - f^\dagger\right\|_{L^2} : \left\|g^{\mathrm{obs}} - Tf^\dagger\right\|_{L^2} \leq \delta\} = \mathcal{O}\left(\log(\delta^{-1})^{-\beta}\right), \qquad \delta \to 0.$$

(iv) $$\left(\inf_{\alpha>0} \mathbf{E}\left[\left\|R_\alpha(Tf^\dagger + \varepsilon W) - f^\dagger\right\|^2_{L^2}\right]\right)^{1/2} = \mathcal{O}\left(\log(\varepsilon^{-1})^{-\beta}\right), \qquad \varepsilon \to 0.$$

Proof. *(i)* $\Leftrightarrow f^\dagger \in \mathbb{X}^T_{\kappa^{2\beta}}$: By Theorem 6.3 we have $f^\dagger \in B^{2\beta}_{2,\infty}(\mathcal{M})$ if and only if $f^\dagger \in \mathbb{X}^T_{\kappa^{2\beta}}$ with $\kappa(\alpha) = ((1/2\bar{t})\ln(\alpha^{-1}))^{-1/2}$ for $0 < \alpha \leq \Lambda_{\mathrm{BH}}(t_0)$ and any $t_0 > 0$.

$f^\dagger \in \mathbb{X}^T_{\kappa^{2\beta}} \Leftrightarrow$ *(ii)*: This follows from Theorem 3.1 since

$$\psi_{\kappa^{2\beta}}(t) = C\log(t^{-1})^{-2\beta}\left(1 + o(1)\right), \qquad \text{as } t \to 0 \tag{38}$$

as shown in [15]. The 3 is included in the definition of ψ to avoid a singularity at $t = 1$.

$f^\dagger \in \mathbb{X}^T_{\kappa^{2\beta}} \Leftrightarrow$ *(iii)*: Follows from Theorems 3.3 and 4.3.

$f^\dagger \in \mathbb{X}^T_{\kappa^{2\beta}} \Leftrightarrow$ *(iv)*: Use the results of [5, §5.1] to see that (32) is fulfilled for any $\tau > 0$ and apply Remark 5.2 and Theorem 3.3. □

7.3. **Sideways heat equation.** We now consider the heat equation in the interval $[0, 1]$. We may think of $[0, 1]$ as the wall of a furnace where the right boundary 1 is the inaccessible interior side and 0 the accessible outer side. We assume the left boundary is insulated and impose the no-flux boundary condition $\partial_x u(0, t) = 0$. The forward problem reads

$$\begin{aligned} &u_t = u_{xx} && \text{in } [0,1] \times \mathbb{R}, \\ &u(1,t) = f(t), && t \in \mathbb{R}, \\ &u_x(0,t) = 0, && t \in \mathbb{R}. \end{aligned}$$

We will consider the inverse problem to estimate the temperature $f(t) = u(1, t)$ at the inaccessible side from measurements of the temperature $g(t) = u(0, t)$ at the accessible side for all times $t \in \mathbb{R}$. As shown in [10] this fits into the framework (35) if we set

$$\Lambda_{\mathrm{SH}}(\mu) = \left|\cosh\sqrt{i\sqrt{\mu}}\right|^{-2}, \qquad \mathcal{M} = \mathbb{R}.$$

Theorem 7.4. *For spectral regularization methods satisfying Assumption (3.2) and the forward operator $T : L^2(\mathbb{R}) \to L^2(\mathbb{R})$ such that $T^*T = \Lambda_{\mathrm{SH}}(-\Delta)$ of the sideways heat equation the following statements for $\beta > 0$ and $f^\dagger \in L^2(\mathbb{R}) \setminus \{0\}$ are equivalent:*

(i) $f^\dagger \in B^{\beta/2}_{2,\infty}(\mathbb{R})$.

(ii) *$f^\dagger$ satisfies a VSC (3) with index function $\psi(t) = C \log(3+t^{-1})^{-2\beta}$ for some $C > 0$.*

(iii) *For a quasioptimal parameter choice rule α_* and a regularization method for which $\Delta(f^\dagger)$ satisfies (22) we have*

$$\sup\{\left\|R_{\alpha_*(\delta,g^{\mathrm{obs}})}g^{\mathrm{obs}} - f^\dagger\right\|_{L^2} : \left\|g^{\mathrm{obs}} - Tf^\dagger\right\|_{L^2} \le \delta\} = \mathcal{O}\left(\log(\delta^{-1})^{-\beta}\right), \qquad \delta \to 0.$$

(iv) $\left(\inf_{\alpha>0} \mathbf{E}\left[\left\|R_\alpha(Tf^\dagger + \varepsilon W) - f^\dagger\right\|_{L^2}^2\right]\right)^{1/2} = \mathcal{O}\left(\log(\varepsilon^{-1})^{-\beta}\right), \qquad \varepsilon \to 0.$

Proof. *(i)* $\Leftrightarrow$ $f^\dagger \in \mathbb{X}^T_{\kappa^{\beta/2}}$*:* As shown in [10] $\Lambda_{\mathrm{SH}}(\mu) = (1/4)\exp(-\sqrt{2}\mu^{1/4})(1 + o(\mu))$ as $\mu \to \infty$. Therefore we obtain $\kappa(\alpha) = 2\ln(\alpha^{-1})^{-2}(1 + o(\alpha))$ as $\alpha \to 0$.

$f^\dagger \in \mathbb{X}^T_{\kappa^{\beta/2}}$ $\Leftrightarrow$ *(ii)* $\Leftrightarrow$ *(iii)* $\Leftrightarrow$ *(iv):* This follows as in proof of Theorem 7.3. Due to the different exponent in the asymptotic formula for κ we have $\psi_{\kappa^{\beta/2}}(t) = C\log(t^{-1})^{-2\beta}(1 + o(1))$ here instead of (38). □

7.4. **Satellite gradiometry.** Let us assume that the Earth is a perfect ball of radius 1. The gravitational potential u of the Earth is determined by its values f at the surface by the exterior boundary value problem

$$\begin{aligned} &\Delta u = 0 && \text{in } \{x \in \mathbb{R}^3 : |x| > 1\} \\ &|u| \to 0, && |x| \to \infty \\ &u = f && \text{on } \mathbb{S}^2 \end{aligned}$$

In satellite gradiometry one studies the inverse problem to determine f from satellite measurements of the rate of change of the gravitational force in radial direction at height $R > 0$, i.e. the data are described by the function $g = \frac{\partial^2 u}{\partial r^2}|_{R\mathbb{S}^2}$. As shown in [10] this fits into the framework (35) if we set

$$\Lambda_{\mathrm{SG}}(\mu) := \left(\frac{1}{2} + \lambda\right)^2 \left(\frac{3}{2} + \lambda\right)^2 R^{-2\lambda}, \qquad \lambda = \sqrt{\frac{1}{2} + \mu}, \qquad \mathcal{M} = \mathbb{S}^2.$$

Note that Λ_{SG} (unlike Λ_{BH} and Λ_{SH}) is not globally monotonically decreasing unless R is large enough (one needs $R \ge \exp((4\sqrt{2} + 2)/(\sqrt{2} + 5)) \approx 3.3$, which is not realistic).

Theorem 7.5. *For spectral regularization methods satisfying Assumption 3.2 and the forward operator $T : L^2(\mathbb{S}^2) \to L^2(\mathbb{S}^2)$ such that $T^*T = \Lambda_{\mathrm{SG}}(-\Delta)$ with R large enough such that Λ_{SG} fulfills Assumption 6.2 of the satellite gradiometry problem the following statements for $\beta > 0$ and $f^\dagger \in L^2(\mathbb{S}^2) \setminus \{0\}$ are equivalent:*

(i) $f^\dagger \in B^\beta_{2,\infty}(\mathbb{S}^2)$.
(ii) *$f^\dagger$ satisfies a VSC (3) with index function $\psi(t) = C\log(3+t^{-1})^{-2\beta}$ for some $C > 0$.*
(iii) *For a quasioptimal parameter choice rule α_* and a regularization method for which $\Delta(f^\dagger)$ satisfies (22) we have*

$$\sup\{\left\|R_{\alpha_*(\delta,g^{\mathrm{obs}})}g^{\mathrm{obs}} - f^\dagger\right\|_{L^2} : \left\|g^{\mathrm{obs}} - Tf^\dagger\right\|_{L^2} \le \delta\} = \mathcal{O}\left(\log(\delta^{-1})^{-\beta}\right), \qquad \delta \to 0.$$

(iv) $\left(\inf_{\alpha>0} \mathbf{E}\left[\left\|R_\alpha(Tf^\dagger + \varepsilon W) - f^\dagger\right\|_{L^2}^2\right]\right)^{1/2} = \mathcal{O}\left(\log(\varepsilon^{-1})^{-\beta}\right), \qquad \varepsilon \to 0.$

Proof. *(i)* $\Leftrightarrow f^\dagger \in \mathbb{X}^T_\kappa$*:* Theorem 6.3 shows that $f^\dagger \in B^\beta_{2,\infty}(\mathbb{S}^2)$ if and only if $f^\dagger \in \mathbb{X}^T_\kappa$ where $\kappa(\alpha) = 2\ln(R)(\ln(\alpha^{-1}))^{-1}(1+o(1))$ as $\alpha \to 0$ since $\Lambda_{\mathrm{SG}}(\mu) = \exp(-2\ln(R)\mu^{1/2})(1+o(1))$ as $\mu \to \infty$.

$f^\dagger \in \mathbb{X}^T_\kappa \Leftrightarrow$ *(ii)* $\Leftrightarrow$ *(iii)* $\Leftrightarrow$ *(iv)*: This follows again along the line of the proof of Theorem 7.3. Here $\psi_{\kappa^\beta}(t) = C\log(t^{-1})^{-2\beta}\,(1+o(1))$ as $t \to 0$. □

8. Conclusions

We have described a general strategy for the verification of VSCs. For linear operators in Hilbert spaces we have shown via a series of equivalence theorems that VSCs are necessary and sufficient for certain rates of convergence both for deterministic errors and for white noise. For a number of relevant inverse problems VSCs with certain index functions are satisfied if and only if the solution belongs to some Besov space.

For other forward operators the set of solutions which satisfies a VSC with a (multiple of a) given index function will not be any known function space. Nevertheless it is interesting to derive verifiable sufficient conditions for VSCs and rates of convergence also for such operators, and we intend to explore the potential of our general strategy in such situations in future research.

Furthermore, our strategy for the verification of VSCs has straightforward extensions to Banach spaces, general data fidelity and penalty functionals, and it has already successfully been applied to nonlinear inverse scattering problems. These extensions will be an interesting topic of future research. Although VSC are known to be sufficient for certain rates of convergence in such general situations, little is known about necessity so far. However, we expect that different techniques than those applied in this paper will be required for such converse results.

Acknowledgement

We would like to thank Uno Hämarik und Toomas Raus for helpful discussions on Lemma 4.2. Financial support by the German Research Foundation DFG through CRC 755, project C09 is gratefully acknowledged.

Appendix A. Spectral source conditions

In this appendix we will use the general strategy of §2 to derive variational source conditions from spectral source conditions. Compared to the implication (ii) $\Rightarrow$ (i) in Theorem 3.1, we can relax the assumption that $t \mapsto \kappa(t)^2/t^{1-\mu}$ is decreasing for some $\mu \in (0,1)$ by allowing also $\mu = 0$. Moreover, the proof for spectral source conditions is considerably simpler.

The result has been known in principle, but previous derivations have been indirect via distance functions, did not yield explicit control over constants, and already for logarithmic source conditions involved quite heavy computations (see [7]).

Proposition A.1. *If T is linear, $\mathbb{Y}$ is a Hilbert space, and $f^\dagger$ satisfies a spectral source condition*

$$f^\dagger = \varphi(T^*T)w, \qquad \|w\| \le \rho$$

with an index function φ such that φ^2 is concave, then $f^\dagger$ satisfies the variational source condition (3) with

$$\psi(\delta^2) = 4\rho^2\varphi\left(\Theta^{-1}\left(\frac{\delta}{\rho}\right)\right)^2, \qquad \Theta(t) := \sqrt{t}\varphi(t). \tag{39}$$

Proof. Let $E_r = 1_{[0,r]}(T^*T)$ denote the spectral family generated by the operator T^*T and set $P_r := I - E_r$ for $r > 0$. Then

$$\|(I - P_r)f^\dagger\|^2 = \|E(r)\varphi(T^*T)w\|^2 = \int_0^r \varphi(t)^2 \,\mathrm{d}\|E_t w\|^2 \le \varphi(r)^2\rho^2.$$

Therefore, (4) holds true with $\kappa(r) = \rho\varphi(r)$. Moreover,

$$\begin{aligned}\left\langle f^\dagger, P_r(f^\dagger - f)\right\rangle &= \left\langle w, P_r\varphi(T^*T)(f^\dagger - f)\right\rangle \le \rho\left(\int_r^\infty \varphi(t)^2 \,\mathrm{d}\|E_t(f^\dagger - f)\|^2\right)^{1/2}\\ &\le \rho\left(\sup_{t\ge r}\frac{\varphi(t)^2}{t}\int_\rho^\infty t\mathrm{d}\|E_t(f^\dagger - f)\|^2\right)^{1/2} \le \rho\frac{\varphi(r)}{\sqrt{r}}\|T(f^\dagger - f)\|\end{aligned}$$

where $\sup_{t\ge r}\varphi(t)^2 t = \varphi(r)^2/r$ since φ^2 is concave and $\varphi(0) = 0$. Hence, (5) holds true with $\sigma(r) = \rho\varphi(r)/\sqrt{r}$ and $C = 0$. Therefore, (3) holds true with

$$\psi(\delta) = 2\inf_{r>0}\left[\rho^2\varphi(r)^2 + \frac{\varphi(r)}{\sqrt{r}}\rho\delta\right] = 2\rho^2\inf_{r>0}\left[\varphi(r)^2 + \frac{\varphi(r)}{\sqrt{r}}\frac{\delta}{\rho}\right].$$

We choose $r = \Theta^{-1}(\delta/\rho)$, i.e. $\sqrt{r}\varphi(r) = \delta/\rho$. Then $\frac{\varphi(r)}{\sqrt{r}}\frac{\delta}{\rho} = \varphi(r)^2 = \varphi(\Theta^{-1}(\delta/\rho))^2$, so we obtain (39). ∎

References

[1] Vinicius Albani, Peter Elbau, Maarten V. de Hoop, and Otmar Scherzer, *Optimal convergence rates results for linear inverse problems in Hilbert spaces* (November 2015), available at 1511.02950.

[2] Andreas Kirsch, *An Introduction to the Mathematical Theory of Inverse Problems*, Springer, New York, Berlin, Heidelberg, 1996.

[3] Roman Andreev, *Tikhonov and Landweber convergence rates: characterization by interpolation spaces*, Inverse Problems **31** (2015Sep), no. 10, 105007.

[4] Roman Andreev, Peter Elbau, Maarten V. de Hoop, Lingyun Qiu, and Otmar Scherzer, *Generalized Convergence Rates Results for Linear Inverse Problems in Hilbert Spaces*, Numerical Functional Analysis and Optimization **36** (2015), no. 5, 549–566.

[5] Nicolai Bissantz, Thorsten Hohage, Axel Munk, and Frits Ruymgaart, *Convergence rates of general regularization methods for statistical inverse problems and applications*, SIAM J. Numer. Anal. **45** (2007), 2610–2636.

[6] Heinz W. Engl, Martin Hanke, and Andreas Neubauer, *Regularization of inverse problems*, Mathematics and its Applications, vol. 375, Kluwer Academic Publishers Group, Dordrecht, 1996.

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

[8] Jens Flemming, Bernd Hofmann, and Peter Mathé, *Sharp converse results for the regularization error using distance functions*, Inverse Problems **27** (2011), no. 2, 025006, 18.
[9] Bernd Hofmann, Barbara Kaltenbacher, Christiane Pöschl, and Otmar Scherzer, *A convergence rates result for Tikhonov regularization in Banach spaces with non-smooth operators*, Inverse Problems **23** (2007), no. 3, 987–1010.
[10] Thorsten Hohage, *Regularization of Exponentially ill-posed Problems*, Numer. Funct. Anal. Optim. **21** (2000), 439–464.
[11] Thorsten Hohage and Frederic Weidling, *Verification of a variational source condition for acoustic inverse medium scattering problems*, Inverse Problems **31** (2015), no. 7, 075006, 14.
[12] Gérard Kerkyacharian and Dominique Picard, *Density estimation by kernel and wavelets methods: optimality of Besov spaces*, Statist. Probab. Lett. **18** (1993), no. 4, 327–336.
[13] ——, *Thresholding algorithms, maxisets and well-concentrated bases*, Test **9** (2000), no. 2, 283–344.
[14] ——, *Minimax or maxisets?*, Bernoulli **8** (2002), no. 2, 219–253.
[15] Bernard A. Mair, *Tikhonov regularization for finitely and infinitely smoothing operators*, SIAM J. Math. Anal. **25** (1994), 135–147.
[16] Peter Mathé and Sergei V. Pereverzev, *Geometry of ill-posed problems in variable Hilbert scales*, Inverse Problems **19** (2003), 789–803.
[17] Andreas Neubauer, *On converse and saturation results for Tikhonov regularization of linear ill-posed problems*, SIAM journal on numerical analysis **34** (1997), no. 2, 517–527.
[18] Paul R Halmos, *What does the spectral theorem say?*, Amer. Math. Monthly **70** (1963), 241–247.
[19] Toomas Raus, *The principle of the residual in the solution of ill-posed problems with nonselfadjoint operator*, Tartu Riikl. Ül. Toimetised **715** (1985), 12–20.
[20] Toomas Raus and Uno Hämarik, *On the quasioptimal regularization parameter choices for solving ill-posed problems*, J. Inverse Ill-Posed Probl. **15** (2007), no. 4, 419–439.
[21] Vincent Rivoirard, *Maxisets for linear procedures*, Statist. Probab. Lett. **67** (2004), no. 3, 267–275.
[22] Hans Triebel, *Theory of Function Spaces II*, Birkhäuser, Basel, 1992.
[23] Frederic Weidling and Thorsten Hohage, *Variational source conditions and stability estimates for inverse electromagnetic medium scattering problems* (December 2015), available at 1512.06586.