Electronic Transactions on Numerical Analysis.
Volume 38, pp. 233-257, 2011.

ISSN 1068-9613.

ETNA
Kent State University
http://etna.math.kent.edu

CONVERGENCE ANALYSIS OF MINIMIZATION-BASED NOISE LEVEL-FREE PARAMETER CHOICE RULES FOR LINEAR ILL-POSED PROBLEMS*

STEFAN KINDERMANN†

Abstract. Minimization-based noise level-free parameter choice rules for the selection of the regularization parameter in linear ill-posed problems are studied. Abstract convergence results for spectral filter regularization operators using a qualitative condition on the (deterministic) data noise are proven. Furthermore, under source conditions on the exact solution, suboptimal convergence rates and, under certain additional regularity conditions, optimal order convergence rates are shown. The abstract results are examined in more detail for several known parameter choice rules: the quasi-optimality rules (both continuous and discrete) and the Hanke-Raus-rules, together with some specific regularization methods: Tikhonov regularization, Landweber iteration, and spectral cutoff.

Key words. regularization, heuristic parameter choice rule, Hanke-Raus rule, quasi-optimality rule, L-curve method

AMS subject classifications. 65J20, 47A52, 65J22

1. Introduction. The proper choice of the regularization parameter in regularization methods is one of the most crucial parts of solving ill-posed problems. Several well-known methods for this task are known. The standard approach is to select the regularization parameter depending on the noise level via a priori or a posteriori rules, usually leading to optimal order convergence rates; see, e.g., [6]. However, these rules need the knowledge (or at least a good guess) of the norm of the noise in the data (noise level). In many cases, the information on the noise level is not available; thus, from a practical point of view, methods which do not make use of the noise level (so called *noise level-free* or *heuristic parameter choice rules*) seem to be most desirable. On the other hand, it has been known for a long time that for ill-posed problems a parameter choice rule that does not depend on the noise level cannot converge in a worst case scenario [1]. Here, a regularization method converges in the worst case if the regularized solution converges to the true solution for *all noisy data* as the noise level tends to 0. The negative result of [1], which sometimes is referred to as the *Bakushinskii veto*, is a strong argument against noise level-free parameter choice rules. Nevertheless, such parameter choice rules are used quite frequently in application and simulation, often yielding reasonable results leaving an unsettling discrepancy between theory and practice.

Recently, [2, 14, 20], a detailed analysis of the quasi-optimality rule, which is a well-known example of a noise level-free rule, revealed that despite the results of Bakushinskii, a convergence analysis is possible if only *restricted noise* is allowed. An appropriate formulation of the restriction on the noise (the noise condition) is a central part of this theory, and was first established in [14]. This result can explain the success of noise level-free rules, because in many practical situation, the data noise does satisfy the noise condition and hence, the regularization method converges, even though, by the Bakushinskii veto, one can always find (or construct) cases in which convergence fails. The analysis can be driven further such that, under the noise condition and smoothness conditions, (in general only suboptimal) convergence rates can be established; see [14, 20].

In this paper, we extend the results of [14, 20] to general minimization-based noise level-free parameter choice rules and general spectral filter-based regularization operators. In this situation, the regularization parameter is selected by minimizing a functional depending on

*Received March 23, 2011. Accepted June 22, 2011. Published online on September 14, 2011. Recommended by L. Reichel.

†Industrial Mathematics Institute, Johannes Kepler University of Linz, Altenbergerstr. 69, 4040 Linz, Austria (kindermann@indmath.uni-linz.ac.at).

ETNA
Kent State University
http://etna.math.kent.edu

the regularization parameter and the given data, but not on the noise level. Let us mention that an important source of ideas for the proofs for this article is [20], where the quasi-optimality rule was analyzed for general linear regularization methods. In this paper, we establish the convergence and convergence rate results of [20] for other parameter choice rules. As particular examples, we study in detail the Hanke-Raus rules, the quasi-optimality rules (continuous and discrete) and to a lesser extent the L-curve method.

The paper is organized as follows. In Section 2, we state some basic definitions and conditions for general spectral filter-based regularization operators and define the minimization-based parameter choice rules in an abstract setting under some standard assumptions, which we impose on the corresponding functionals.

In Section 3, we prove in a general framework (i.e., for general regularization operators and general parameter choice functionals) convergence and convergence rates for these methods.

In Section 4, we define some specific examples of noise level-free parameter choice rules and apply the convergence result of Section 3 to these cases. The conditions there are stated for general spectral filter-based regularizations.

In Section 5, we verify these conditions for some prototypical regularization methods, namely Tikhonov regularization, Landweber iteration and spectral cutoff in connection with the above mentioned parameter choice rules. Furthermore, in this section, we explain the drawback of the L-curve method.

Finally, in Section 6, we review the results and interpret the stated conditions on an informal level.

2. Noise level-free parameter choice rules. We consider linear ill-posed equations in Hilbert spaces,

$$Ax = y,$$

where $A : X \to Y$ is an operator between Hilbert spaces X, Y, with x the unknown solution and y the given data. Such equations can be approximately solved by regularization operators. In the following, we study linear spectral filter-based regularization operators. Suppose that we are given a family of regularization operators

$$R_\alpha : Y \to X, \quad \alpha \in M \subset (0, \alpha_0],$$

with M being a set of possible regularization parameters such that

$$\overline{M} = M \cup \{0\}, \tag{2.1}$$

where $\overline{M}$ denotes the closure of M. In particular, by this condition, 0 is a limit point of M. We consider regularization operators defined by spectral filter functions that satisfy the general conditions of a regularization method [6]. More precisely, we impose the following:

DEFINITION 2.1. *Let M satisfy (2.1). A spectral filter is a family of piecewise continuous functions $g_\alpha : [0, \|A\|^2] \to \mathbb{R}$, $\alpha \in M$ satisfying*

- *there exists a constant C_g and for all $\tau > 0$ a constant G_τ with*

$$\sup_{\alpha \in \overline{M}} \sup_{\lambda \in [0, \|A\|^2]} |\lambda g_\alpha(\lambda)| \leq C_g, \tag{2.2}$$

$$\sup_{\alpha \in M \cap [\tau, \alpha_0]} \sup_{\lambda \in [0, \|A\|^2]} |g_\alpha(\lambda)| \leq G_\tau, \tag{2.3}$$

- *for all $\lambda \in (0, \|A\|^2]$*

$$\lim_{M \ni \alpha \to 0} g_\alpha(\lambda) = \frac{1}{\lambda}.$$

ETNA
Kent State University
http://etna.math.kent.edu

For a family of spectral filter functions we define the residual functions

$$r_\alpha(\lambda) := 1 - \lambda g_\alpha(\lambda).$$

In all of the following we consider only regularization operators defined by a spectral filter, i.e,

$$R_\alpha y = \int g_\alpha(\lambda) dE_\lambda \, A^* y.$$

Using the notation of [6], E_λ denotes a spectral family of A^*A, a spectral family of AA^* will be denoted by F_λ, and Q denotes the orthogonal projector onto $\overline{R(A)}$.

A convergence rate analysis will be derived for monotone spectral filters.

DEFINITION 2.2. *We say that a function g_α is a monotone spectral filter if for all $\lambda \in (0, \|A\|^2]$*

$$M \ni \alpha \mapsto |g_\alpha(\lambda)| \text{ is monotonically decreasing,} \tag{2.4}$$

$$M \ni \alpha \mapsto |r_\alpha(\lambda)| \text{ is monotonically increasing}. \tag{2.5}$$

An *index function* [19] is a function $\phi : \mathbb{R}^+ \to \mathbb{R}^+$ that is continuous and strictly monotonically increasing and satisfies $\phi(0) = 0$. For the convergence rate analysis we will need some further conditions on the filter functions.

- There is a constant $\overline{C}_c$ such that

$$|g_\alpha(\lambda)| \leq \frac{\overline{C}_c}{\alpha} \quad \forall\, 0 \leq \lambda \leq \|A\|^2. \tag{2.6}$$

- There is a constant $0 < \eta \leq \|A\|^2$ such that for all $0 < \gamma \leq 1$ there exists a constant $\underline{C}_{l,\gamma}$ such that

$$|g_\alpha(\lambda)| \geq \frac{\underline{C}_{l,\gamma}}{\alpha} \quad \forall\, 0 \leq \lambda \leq \gamma\alpha \quad 0 \leq \gamma\alpha \leq \eta, \quad \alpha \in M. \tag{2.7}$$

- There is a constant $0 < \eta \leq \|A\|^2$ such that for all $0 < \gamma \leq 1$ there exists a constant $\underline{C}_{h,\gamma}$ such that

$$|g_\alpha(\lambda)| \geq \frac{\underline{C}_{h,\gamma}}{\lambda} \quad \forall\, \gamma\alpha \leq \lambda \leq \|A\|^2 \quad 0 \leq \gamma\alpha \leq \eta, \quad \alpha \in M. \tag{2.8}$$

- There is a constant $0 < \eta \leq \|A\|^2$ such that for all $0 < \gamma \leq 1$ there exists a constant $\underline{D}_{l,\gamma}$ such that

$$|r_\alpha(\lambda)| \geq \underline{D}_{l,\gamma} \quad \forall\, 0 < \lambda \leq \gamma\alpha \quad 0 \leq \gamma\alpha \leq \eta, \quad \alpha \in M. \tag{2.9}$$

Similar conditions were used in [20]. We also need the concept of qualification; see, e.g., [6] and the generalization in [19].

DEFINITION 2.3. *We say that an index function ρ has a qualification (for the spectral filter g_α) if there is a constant D_ρ such that*

$$|r_\alpha(\lambda)|\rho(\lambda) \leq D_\rho \rho(\alpha) \quad \forall \alpha \in M, \lambda \in (0, \|A\|^2]. \tag{2.10}$$

We say that $\mu_0 \in \mathbb{R}^+$ is a qualification index if $\rho(\lambda) = \lambda^{\mu_0}$ is a qualification and there is a constant $0 < \eta < \|A\|^2$ such that for all $0 < \gamma \leq 1$ there exists a constant $D_{ind,\mu_0,\gamma}$ such that

$$|r_\alpha(\lambda)|\lambda^{\mu_0} \geq D_{ind,\mu_0,\gamma}\alpha^{\mu_0} \quad \forall\, \gamma\alpha \leq \lambda \leq \|A\|^2 \quad 0 \leq \gamma\alpha \leq \eta, \quad \alpha \in M. \tag{2.11}$$

ETNA
Kent State University
http://etna.math.kent.edu

Note that the notion of qualification is sometimes used differently. Many authors simply refer to the qualification index μ_0 in Definition 2.3 as the qualification. In this paper, the qualification is an index function as in [19], and to distinguish it from the classical qualification μ_0 we refer to this number as the qualification index.

For a continuous regularization method, the regularization parameter is usually chosen in some interval

$$M = M_c = (0, \alpha_0),$$

while for a discrete regularization method the regularization parameter (usually the inverse of an iteration index) is in a discrete set

$$M = M_d = \bigcup_{i=1}^{\infty} \{\alpha_i\} \quad \alpha_i \text{ strictly monotonically decreasing with } \lim_{i\to\infty} \alpha_i = 0.$$

Let us give some examples of regularization operators:

- Tikhonov regularization, $M_c = (0, \alpha_0)$

$$g_\alpha(\lambda) = \frac{1}{\lambda + \alpha}, \quad r_\alpha(\lambda) = \frac{\alpha}{\lambda + \alpha},$$

- Landweber iteration, $\|A\| \leq 1$, $M_d = \{\frac{1}{k}, k \in N\}$, $\alpha = \frac{1}{k}$,

$$g_\alpha(\lambda) = g_{\frac{1}{k}}(\lambda) = \sum_{i=0}^{k-1} (1-\lambda)^i, \quad r_{\frac{1}{k}}(\lambda) = (1-\lambda)^k,$$

- spectral cutoff: $M_c = (0, \alpha_0)$ (continuous) or $M = M_d = \{\sigma_i\}$ (truncated singular value decomposition), where the σ_i are the singular values of the compact operator A.

$$g_\alpha(\lambda) = \begin{cases} \frac{1}{\lambda} & \lambda \geq \alpha, \\ 0 & \lambda < \alpha, \end{cases} \qquad r_\alpha(\lambda) = \begin{cases} 0 & \lambda \geq \alpha, \\ 1 & \lambda < \alpha. \end{cases}$$

Note that any continuous regularization method can be made a discrete one by restricting the set of allowed regularization parameters to a discrete set.

All these regularization operators are defined by monotone spectral filter functions. Moreover, Tikhonov regularization and Landweber iteration satisfy (2.6)–(2.9); spectral cutoff satisfies (2.6), (2.8), (2.9), while (2.7) does not hold; cf. [20].

Let us furthermore introduce some standard notation. We will denote by $y \in D(A^\dagger)$ the (unknown) exact data, and by $x^\dagger = A^\dagger y$ the unknown exact solution. In practice, only a noisy version of y is known, which we denote by y_δ. For a given y_δ we define the noise level

$$\|y_\delta - y\|_Y = \delta \quad y \in D(A^\dagger)$$

in the usual way. As already mentioned, for noise level-free parameter choice rules, knowledge of δ is *not* used.

For the heuristic rules in this paper the regularization parameter is chosen as a minimizer of a functional $\alpha \mapsto \psi(\alpha, y_\delta)$. Let R_α be a fixed family of regularization operators with M as in (2.1). We consider rules using certain positive functionals ψ that satisfy the following conditions.

ETNA
Kent State University
http://etna.math.kent.edu

ASSUMPTION 2.4.

A1. ψ is nonnegative:

$$\psi : M \times Y \to \mathbb{R}_0^+.$$

A2. For all $\alpha \in M$, $y \in Y$, ψ is symmetric:

$$\psi(\alpha, -y) = \psi(\alpha, y).$$

A3. For any $\alpha \in M$,

$$\psi(\alpha, .) : Y \to \mathbb{R}_0^+$$

is continuous.

A4. For any $z \in Y$,

$$\psi(., z) : M \to \mathbb{R}_0^+$$

is lower semicontinuous.

A5. If $z \in D(A^\dagger)$, then

$$\lim_{\alpha \to 0} \psi(\alpha, z) = 0.$$

It will be shown that most of the well-known minimization-based noise level-free parameter choice rules correspond to a functional ψ that satisfies these conditions.

We now state a class of parameter choice rules: Given a functional ψ that satisfies Assumption 2.4, we define a regularization parameter $\alpha^*(y_\delta)$ as

$$\alpha^*(y_\delta) := \begin{cases} \text{argmin}_{\alpha \in M} \psi(\alpha, y_\delta) & \text{if a minimum in } M \text{ exists,} \\ 0 & \text{else.} \end{cases} \tag{2.12}$$

In the case that there are multiple global minima, we simply select an arbitrary one; the convergence properties will not depend on the specific choice. Obviously, we do not need any information on the noise level to choose $\alpha^*(y_\delta)$, but only the given noisy data y_δ.

It is convenient to extend the definition of ψ to $\alpha = 0$:

$$\begin{aligned} \overline{\psi}(\alpha, y_\delta) : \overline{M} \times Y &\to [0, \infty] \\ \overline{\psi}(\alpha, y_\delta) &:= \begin{cases} \psi(\alpha, y_\delta) & \text{if } \alpha > 0 \\ \liminf_{\tau \to 0} \psi(\tau, y_\delta) & \text{if } \alpha = 0. \end{cases} \end{aligned} \tag{2.13}$$

Note that a realization of the parameter choice $\alpha^*(y_\delta)$ can also be written as

$$\alpha^*(y_\delta) = \max\{\text{argmin}_{\alpha \in \overline{M}} \overline{\psi}(\alpha, y_\delta)\}.$$

3. Convergence and convergence rates. We analyze the convergence of spectral filter-based regularization methods with the parameter choice rules defined in (2.12). At first, we study conditions that yield convergence of such methods. We remind the reader that δ denotes the noise level, y_δ are the given noisy data, and y denotes the exact data.

ETNA
Kent State University
http://etna.math.kent.edu

3.1. Abstract convergence result. The main result in this subsection is Theorem 3.6.

PROPOSITION 3.1. *Let Assumption 2.4 hold, and let the parameter $\alpha^*(y_\delta)$ be defined by (2.12). Then*

$$\lim_{\delta\to 0} \overline{\psi}(\alpha^*(y_\delta), y_\delta) = 0.$$

Proof. Let $y \in D(A^\dagger)$ be fixed and y_{δ_k} be a sequence of noisy data such that their noise levels satisfy $\delta_k \to 0$. By definition, for arbitrary $\alpha \in M$ fixed, we have

$$\overline{\psi}(\alpha^*(y_\delta), y_{\delta_k}) \leq \psi(\alpha, y_{\delta_k}).$$

By continuity (A3) we have that $\lim_{\delta_k\to 0} \psi(\alpha, y_{\delta_k}) = \psi(\alpha, y)$; hence for all $\alpha \in M$,

$$\limsup_{\delta_k\to 0} \overline{\psi}(\alpha^*(y_\delta), y_{\delta_k}) \leq \psi(\alpha, y).$$

According to (A5) the right-hand side in this inequality tends to 0 as $\alpha \to 0$; hence

$$0 \leq \liminf_{\delta_k\to 0} \overline{\psi}(\alpha^*(y_\delta), y_{\delta_k}) \leq \limsup_{\delta_k\to 0} \overline{\psi}(\alpha^*(y_\delta), y_{\delta_k}) \leq 0,$$

which proves the proposition. □

In order to prove convergence, one has to impose additional conditions. The first one is a consistency condition that relates ψ to the approximation error.

CONDITION 3.2 (Consistency). *Let $z \in D(A^\dagger)$ be fixed. For all $\alpha \in M$, and all sequences $(\alpha_n)_n \in M$, $(z_n)_n \in Y$ with $\lim_{n\to\infty} \alpha_n = \alpha$ and $\lim_{k\to\infty} z_n \to z$ it holds that*

$$\lim_{n\to\infty} \psi(\alpha_n, z_n) = 0 \Rightarrow \lim_{n\to\infty} \|R_{\alpha_n} z - A^\dagger z\| = 0.$$

Second, we need a noise condition, which is the most important part in the convergence theory of noise level-free parameter choice rules. This condition has to take into account that a uniform convergence proof for noise level-free parameter choice rules is impossible.

CONDITION 3.3 (Noise condition). *There exists a set $\mathcal{N}_z \subset Y$ such that for all $z_n \in \mathcal{N}_z$ with $\lim_{n\to\infty} z_n = z$*

$$\overline{\psi}(0, z_n) > \inf_{\alpha\in M} \psi(\alpha, z_n), \tag{3.1}$$

and for all $\alpha_n \in M$ with $\lim_{n\to\infty} \alpha_n = 0$

$$\lim_{n\to\infty} \psi(\alpha_n, z_n) = 0 \Rightarrow \lim_{n\to\infty} \|R_{\alpha_n} z_n - R_{\alpha_n} z\| \to 0. \tag{3.2}$$

We notice that (3.1) can only hold if $\mathcal{N}_z \cap D(A^\dagger) = \emptyset$ according to (A5) (if we exclude the degenerate case $\psi(\alpha, z) = 0, \forall \alpha \in M$). The noise condition is a central part of the analysis in this paper. Let us again emphasize the difference from the conventional worst case convergence analysis: there, convergence is analyzed by treating the following error for some parameter choice rule $\alpha^*(y_\delta)$

$$\lim_{\delta\to 0} \|R_{\alpha^*(y_\delta)} y_\delta - A^\dagger y\| \quad \forall y_\delta \in Y \text{ with } \|y_\delta - y\| \leq \delta,$$

while in this paper we only consider the error with a noise restriction

$$\lim_{\delta\to 0} \|R_{\alpha^*(y_\delta)} y_\delta - A^\dagger y\| \quad \forall y_\delta \in Y \text{ with } \|y_\delta - y\| \leq \delta \text{ and } y_\delta \in \mathcal{N}_y.$$

ETNA
Kent State University
http://etna.math.kent.edu

While the first limit (worst case) cannot tend to 0 for heuristic parameter choice rules by the Bakushinskii veto, we will show that the second one (restricted noise case) does.

Before we come to the convergence theorem, let us discuss some sufficient conditions for Condition 3.2.

LEMMA 3.4. *Let ψ satisfy Assumption 2.4, and for a neighborhood $U(z)$ of $z \in D(A^\dagger)$ let ψ be lower semicontinuous on $M \cap [\tau, \alpha_0) \times U(z)$, for all $\tau > 0$. If additionally for all $\alpha \in M$ it holds that*

$$\psi(\alpha, z) > 0, \tag{3.3}$$

then Condition 3.2 is satisfied for z. Moreover, in this case $\lim_{\delta\to 0} \alpha^*(y_\delta) = 0$.

Proof. With the notation as in Condition 3.2, it follows from the lower semicontinuity of ψ that

$$0 < \psi(\alpha, z) \leq \liminf_{n\to\infty} \psi(\alpha_n, z_n),$$

which contradicts (3.3). Hence, the antecedent in Assumption 2.4 is always false, which makes the implication always true. Now suppose that $\lim_{\delta\to 0} \alpha^*(y_\delta) \neq 0$. Then we can find a subsequence of $\alpha^*(y_\delta)$ such that $\alpha^*(y_{\delta_k}) \to \alpha \neq 0$. With the same argument of lower semicontinuity and Proposition 3.1 it follows that $\psi(\alpha, y) = 0$, which is again a contradiction to (3.3), hence, $\alpha^*(y_\delta) \to 0$. ◻

The positivity of ψ in (3.3) is not always satisfied. In such situations the following lemma is useful.

LEMMA 3.5. *Let ψ satisfy Assumption 2.4, and for a neighborhood $U(z)$ of $z \in D(A^\dagger)$ let ψ be lower semicontinuous on $M \cap [\tau, \alpha_0) \times U(z)$, for all $\tau > 0$. For all $\lambda \in (0, \|A\|^2]$, let the function $\alpha \mapsto r_\alpha(\lambda)$ be upper semicontinuous at any point $\alpha \in M$. If it holds that at $z \in D(A^\dagger)$ and for all $\alpha \in M$*

$$\psi(\alpha, z) = 0 \Rightarrow R_\alpha z = A^\dagger z, \tag{3.4}$$

then Condition 3.2 is satisfied.

Proof. With the notation of Condition 3.2 and from lower semicontinuity we find as in the proof of Lemma 3.4 that $\psi(\alpha, z) = 0$. From the assumptions on $r_\alpha(\lambda)$, (3.4) and Fatou's lemma we get

$$\limsup_{\alpha_n\to\alpha} \|R_{\alpha_n} z - A^\dagger z\|^2 = \limsup_{\alpha_n\to\alpha} \int |r_{\alpha_n}(\lambda)|^2 dE_\lambda \|A^\dagger z\|^2$$
$$\leq \int \limsup_{\alpha_n\to\alpha} |r_{\alpha_n}(\lambda)|^2 dE_\lambda \|A^\dagger z\|^2 \leq \int |r_\alpha(\lambda)|^2 dE_\lambda \|A^\dagger z\|^2 = \|R_\alpha z - A^\dagger z\|^2 = 0.$$

Thus, we arrive at $\lim_{n\to\infty} \|R_{\alpha_n} z - A^\dagger z\|^2 = 0$ which validates Condition 3.2. ◻

We now come to the main convergence proof:

THEOREM 3.6 (Convergence theorem). *Let R_α be a regularization operator defined by a spectral filter, and let ψ satisfy Assumption 2.4. Let Condition 3.2, Condition 3.3 hold and for $z = y = Ax^\dagger$ let $y_\delta \in \mathcal{N}_y$. Then $\alpha^*(y_\delta) > 0$ and*

$$R_{\alpha^*(y_\delta)} y_\delta \to A^\dagger y \quad \text{for } \delta \to 0. \tag{3.5}$$

Proof. With $y = Ax^\dagger$ the following decomposition is standard

$$\|R_{\alpha^*(y_\delta)} y_\delta - A^\dagger y\| \leq \|R_{\alpha^*(y_\delta)} y_\delta - R_{\alpha^*(y_\delta)} y\| + \|R_{\alpha^*(y_\delta)} y - A^\dagger y\|. \tag{3.6}$$

Let y_{δ_k} be a sequence of noisy data with noise level $\delta_k \to 0$ and let $\alpha^*(y_{\delta_{k_j}})$ be an arbitrary subsequence of $\alpha^*(y_{\delta_k})$. Since this sequence is bounded, it has a converging subsequence again denoted by $\alpha^*(y_{\delta_{k_j}})$: $\lim_{j\to\infty} \alpha^*(y_{\delta_{k_j}}) = \alpha$. We distinguish the cases $\alpha = 0$ and $\alpha > 0$.

First, assume that $\alpha = 0$. In this case, it follows from the general convergence theory for regularization methods [6] that

$$\lim_{j\to\infty} \|R_{\alpha^*(y_{\delta_{k_j}})} y - A^\dagger y\| = 0.$$

From Proposition 3.1 we obtain that $\overline{\psi}(\alpha^*(y_{\delta_{k_j}}), y_{\delta k_j}) \to 0$. From (3.1) it follows that $\alpha^*(y_{\delta_{k_j}}) \neq 0$, hence $\overline{\psi}(\alpha^*(y_{\delta_{k_j}}), y_{\delta k_j}) = \psi(\alpha^*(y_{\delta_{k_j}}), y_{\delta k_j})$. By (3.2), we conclude that

$$\lim_{j\to\infty} \|R_{\alpha^*(y_{\delta_{k_j}})} y_{\delta k_j} - R_{\alpha^*(y_{\delta_{k_j}})} y\| = 0.$$

Thus, from (3.6) we obtain that (3.5) holds for the subsequence $(R_{\alpha^*(y_{\delta_{k_j}})} y_{\delta k_j})_j$.

Now assume that $\alpha > 0$. Proposition 3.1 and Condition 3.2 imply that

$$\lim_{j\to\infty} \|R_{\alpha^*(y_{\delta_{k_j}})} y - A^\dagger y\| = 0.$$

Moreover, from (2.2), (2.3), it follows that for j sufficiently large and τ sufficiently small

$$\|R_{\alpha^*(y_{\delta_{k_j}})} y_{\delta k_j} - R_{\alpha^*(y_{\delta_{k_j}})} y\| \leq \sqrt{C_g G_\tau} \|y_{\delta k_j} - y\| \to 0 \quad \text{as } j \to \infty.$$

Together, we can conclude that (3.5) holds for the subsequence $(R_{\alpha^*(y_{\delta_{k_j}})} y_{\delta k_j})_j$. In any case, we have shown that any subsequence of $R_{\alpha^*(y_\delta)} y_\delta$ has a subsequence converging to $A^\dagger y$, thus, (3.5) must hold. □

We note that for Theorem 3.6 it is not necessary that ψ satisfies (A2) of Assumption 2.4.

3.2. Abstract convergence rates. The assumptions on ψ in the previous section are not enough to prove convergence rates. In this section, we study convergence rates for subadditive functionals ψ (i.e., functionals for which Condition 3.7 below is satisfied). Additionally, we have to impose quantitative versions of Condition 3.2 and Condition 3.3.

A major simplification is obtained if we assume that ψ is subadditive:

CONDITION 3.7. *There exists a constant $\kappa_s > 0$, such that for all $\alpha \in M$, $z, e \in Y$,*

$$\psi(\alpha, z + e) \leq \kappa_s (\psi(\alpha, z) + \psi(\alpha, e)).$$

For subadditive and uniformly continuous ψ the following inequality suffices to satisfy Condition 3.2.

CONDITION 3.8. *There exists an index function Φ such that for all $\alpha \in M$, $z \in D(A^\dagger)$:*

$$\|R_\alpha z - A^\dagger z\| \leq \Phi(\psi(\alpha, z)). \tag{3.7}$$

LEMMA 3.9. *Let ψ satisfy Assumption 2.4, Condition 3.7 and suppose that for any $\tau > 0$*

$$\lim_{z_k\to 0} \psi(\alpha, z_k) = 0 \quad \textit{uniformly in } \alpha \in M \cap [\tau, \infty). \tag{3.8}$$

If ψ additionally satisfies Condition 3.8 then Condition 3.2 holds for all $z \in D(A^\dagger)$.

Proof. In the situation of Condition 3.2 we have that

$$\Phi^{-1}\left(\|R_{\alpha_n} z - A^\dagger z\|\right) \leq \psi(\alpha_n, z) \leq \kappa_s (\psi(\alpha_n, z_n) + \psi(\alpha_n, z - z_n)).$$

ETNA
Kent State University
http://etna.math.kent.edu

The first term on the right-hand side tends to 0 by the hypothesis in Condition 3.2 while the second term tends to 0 by uniform continuity (3.8). From the properties of the index function, this means that $\|R_{\alpha_n} z - A^\dagger z\|$ tends to 0. □

A quantitative alternative to Condition 3.3 is the following

CONDITION 3.10. *There exists a set* $\mathcal{N} \subset Y$ *with* $\mathcal{N} \cap D(A^\dagger) = \emptyset$ *and a constant* $\kappa_l > 0$ *such that for all* $z_n - z \in \mathcal{N}$, $\alpha \in M$,

$$\kappa_l \|R_\alpha z_n - R_\alpha z\| \le \psi(\alpha, z_n - z). \tag{3.9}$$

For subadditive ψ, Condition 3.10 is indeed sufficient for Condition 3.3 as the following lemma shows:

LEMMA 3.11. *Let* ψ *satisfy Assumption 2.4 and let Condition 3.7 and Condition 3.10 hold. Then Condition 3.3 holds with* $\mathcal{N}_z = z + \mathcal{N}$. *Moreover, in this case we have that*

$$\lim_{\alpha\to 0} \psi(\alpha, z_n - z) = \lim_{\alpha\to 0} \psi(\alpha, z_n) = \infty. \tag{3.10}$$

Proof. From Condition 3.7, (A5) and in the situation of Condition 3.3 it follows from (3.9) that

$$\begin{aligned} 0 &\le \lim_{n\to\infty} \|R_{\alpha_n} z_n - R_{\alpha_n} z\| \le \frac{1}{\kappa_l} \lim_{n\to\infty} \psi(\alpha_n, z_n - z) \\ &\le \frac{\kappa_s}{\kappa_l} \left(\lim_{n\to\infty} \psi(\alpha_n, z_n) + \lim_{n\to\infty} \psi(\alpha_n, -z) \right) = 0, \end{aligned}$$

which shows (3.2). To prove (3.1), we observe that since $z_n - z \notin D(A^\dagger)$, it holds that $\lim_{\alpha\to 0} \|R_\alpha(z_n - z)\| = \infty$ (cf. [6, Prop. 3.6]). With

$$\kappa_l \|R_\alpha(z_n - z)\| \le \psi(\alpha, z_n - z) \le \kappa_s \left(\psi(\alpha, z_n) + \psi(\alpha, -z)\right)$$

and (A5) we conclude that (3.10) and consequently (3.1) holds. □

We note that the condition (3.9) was already used for the quasi-optimality rule in [7]. However, finding a specific set $\mathcal{N}$ and a constant κ_l such that (3.9) can be verified is a difficult task. Such estimates with specific sets $\mathcal{N}$ were first established in [14] for the quasi-optimality rule. We will show in Section 4.1 that similar sets $\mathcal{N}$ as in [14] can also be used for other parameter choice rules.

We now come to the main abstract convergence rate result:

THEOREM 3.12 (Convergence rate theorem). *Let* g_α *be a monotone spectral filter (i.e.,* (2.4), (2.5) *holds). Let* ψ *satisfy Assumption 2.4, Condition 3.7, Condition 3.10 and Condition 3.8. Moreover, for any* $z \in D(A^\dagger)$ *and* $e \in \mathcal{N}_z$, *let there exists a monotonically increasing function* $\rho_{\uparrow,z}(\alpha)$ *and a monotonically decreasing function* $\rho_{\downarrow,e}(\alpha)$ *such that*

$$\psi(\alpha, z) \le \rho_{\uparrow,z}(\alpha) \quad \forall \alpha \in M, \tag{3.11}$$

$$\psi(\alpha, e) \le \rho_{\downarrow,e}(\alpha) \quad \forall \alpha \in M. \tag{3.12}$$

If $y_\delta \in \mathcal{N}_y$, *then there exists a constant* $\tilde{C}$ *such that*

$$\begin{aligned} \|R_{\alpha^*(y_\delta)} y^\delta - A^\dagger y\| \le \inf_{\alpha>0} \Big\{ &\Phi \left[\tilde{C} \left(\rho_{\uparrow,y}(\alpha) + \rho_{\downarrow,y-y_\delta}(\alpha) \right) \right] \\ &+ \tilde{C} \left(\rho_{\uparrow,y}(\alpha) + \rho_{\downarrow,y-y_\delta}(\alpha) \right) \Big\}. \end{aligned} \tag{3.13}$$

Proof. Let $M \ni \bar{\alpha} > 0$ be arbitrary but fixed. From (2.4) it follows that the propagated data error $\alpha \mapsto \|R_\alpha(y_\delta - y)\|$ is monotonically decreasing and that for all $y \in D(A^\dagger)$ the approximation error $\alpha \to \|R_\alpha y - A^\dagger y\|$ is monotonically increasing. Suppose that $\bar{\alpha} \geq \alpha^*(y_\delta)$. Then, by monotonicity we have

$$\|R_{\alpha^*(y_\delta)}y - A^\dagger y\| \leq \|R_{\bar{\alpha}}y - A^\dagger y\| \leq \Phi\left(\psi(\bar{\alpha}, y)\right) \leq \Phi\left(\rho_{\uparrow,y}(\bar{\alpha})\right).$$

Moreover, Condition 3.10 and Condition 3.7 imply

$$\begin{aligned}\kappa_l\|R_{\alpha^*(y_\delta)}(y_\delta - y)\| &\leq \psi(\alpha^*(y_\delta), y_\delta - y) \leq \kappa_s\psi(\alpha^*(y_\delta), y_\delta) + \kappa_s\psi(\alpha^*(y_\delta), -y)\\ &\leq \kappa_s\psi(\bar{\alpha}, y_\delta) + \kappa_s\psi(\alpha^*(y_\delta), -y) \leq \kappa_s^2\psi(\bar{\alpha}, y_\delta - y) + \kappa_s^2\psi(\bar{\alpha}, y) + \kappa_s\rho_{\uparrow,y}(\alpha^*(y_\delta))\\ &\leq (\kappa_s^2 + \kappa_s)\rho_{\uparrow,y}(\bar{\alpha}) + \kappa_s^2\rho_{\downarrow,y_\delta - y}(\bar{\alpha}).\end{aligned}$$

Combining these bounds we obtain a constant C such that

$$\|R_{\alpha^*(y_\delta)}y^\delta - A^\dagger y\| \leq \Phi\left(\rho_{\uparrow,y}(\bar{\alpha})\right) + C\left(\rho_{\uparrow,y}(\bar{\alpha}) + \rho_{\downarrow,y_\delta - y}(\bar{\alpha})\right).$$

On the other hand, for $\bar{\alpha} \leq \alpha^*(y_\delta)$, we see that

$$\|R_{\alpha^*(y_\delta)}(y_\delta - y)\| \leq \|R_{\bar{\alpha}}(y_\delta - y)\| \leq \frac{1}{\kappa_l}\rho_{\downarrow,y_\delta - y}(\bar{\alpha}),$$

and

$$\begin{aligned}\Phi^{-1}\left(\|R_{\alpha^*(y_\delta)}y - A^\dagger y)\|\right) &\leq \psi(\alpha^*(y_\delta), y)\\ &\leq \kappa_s\psi(\alpha^*(y_\delta), y - y_\delta) + \kappa_s\psi(\alpha^*(y_\delta), y_\delta)\\ &\leq \kappa_s\psi(\alpha^*(y_\delta), y - y_\delta) + \kappa_s\psi(\bar{\alpha}, y_\delta)\\ &\leq \kappa_s\psi(\alpha^*(y_\delta), y - y_\delta) + \kappa_s^2\psi(\bar{\alpha}, y_\delta - y) + \kappa_s^2\psi(\bar{\alpha}, y)\\ &\leq \kappa_s^2\rho_{\uparrow,y}(\bar{\alpha}) + (\kappa_s^2 + \kappa_s)\rho_{\downarrow,y_\delta - y}(\bar{\alpha}).\end{aligned}$$

Thus, in this case, we obtain that

$$\|R_{\alpha^*(y_\delta)}y^\delta - A^\dagger y\| \leq \Phi\left(C\left(\rho_{\uparrow,y}(\bar{\alpha}) + \rho_{\downarrow,y_\delta - y}(\bar{\alpha})\right)\right) + \frac{1}{\kappa_l}\rho_{\downarrow,y_\delta - y}(\bar{\alpha}).$$

In either case we have shown (3.13). □

As a special case of the previous theorem we obtain order optimal estimates:

THEOREM 3.13 (Optimal order). *Let g_α be a monotone spectral filter. Let ψ satisfy Assumption 2.4 and Condition 3.7. Moreover, for any $z \in D(A^\dagger)$ and $e \in \mathcal{N}_z$, let the following inequalities hold with some positive constants $\kappa_1, \kappa_2, \kappa_3, \kappa_4$*

$$\kappa_1\|R_\alpha z - A^\dagger z\| \leq \psi(\alpha, z) \leq \kappa_3\|R_\alpha z - A^\dagger z\| \quad \forall\alpha \in M, \tag{3.14}$$

$$\kappa_2\|R_\alpha e\| \leq \psi(\alpha, e) \leq \kappa_4\|R_\alpha e\| \quad \forall\alpha \in M. \tag{3.15}$$

If $y_\delta \in \mathcal{N}_y$, then there exists a constant C such that

$$\|R_{\alpha^*(y_\delta)}y^\delta - A^\dagger y\| \leq C\inf_{\alpha>0}\left(\|R_\alpha y - A^\dagger y\| + \|R_\alpha(y - y_\delta)\|\right).$$

This theorem gives an oracle type estimate, since the total error of the regularization is of the order of the error for the best possible choice of the regularization parameter.

If subadditivity does not hold we still can prove the following result.

ETNA
Kent State University
http://etna.math.kent.edu

THEOREM 3.14 (Convergence rate without subadditivity). *Let g_α be a monotone spectral filter. Let ψ satisfy Assumption 2.4 and suppose that for any $z \in D(A^\dagger)$ and $e \in \mathcal{N}_z$ there exists a monotonically increasing function $f_{\uparrow,z}(\alpha)$, and a monotonically decreasing function $f_{\downarrow,e}(\alpha)$ and there exist an index function Φ such that for all $z_n - z \in \mathcal{N}_z$,*

$$\begin{aligned} \|R_\alpha z - A^\dagger z\| &\leq \Phi\left(\psi(\alpha, z_n) + f_{\downarrow, z-z_n}(\alpha)\right), \\ \tilde{C}\|R_\alpha z_n - R_\alpha z\| &\leq \psi(\alpha, z_n) + f_{\uparrow,z}(\alpha). \end{aligned} \tag{3.16}$$

Then there is a constant

$$\begin{aligned} \|R_{\alpha^*(y_\delta)} y^\delta - A^\dagger y\| \leq \inf_{\alpha > 0} \Big\{ &\Phi\left[\psi(\alpha, y_\delta) + f_{\downarrow, y-y_\delta}(\alpha)\right] \\ &+ \tilde{C}\left(\psi(\alpha, y_\delta) + f_{\uparrow,y}(\alpha)\right) + f_{\downarrow, y-y_\delta}(\alpha) \Big\}. \end{aligned}$$

Proof. As before, let $\bar{\alpha}$ be arbitrary. Then if $\alpha^*(y_\delta) \leq \bar{\alpha}$ we obtain the inequalities

$$\|R_{\alpha^*(y_\delta)} y - A^\dagger y\| \leq \|R_{\bar{\alpha}} y - A^\dagger y\| \leq \Phi\left(\psi(\bar{\alpha}, y_\delta) + f_{\downarrow, y-y_\delta}(\bar{\alpha})\right),$$

and

$$\begin{aligned} \|R_{\alpha^*(y_\delta)} y_\delta - R_{\alpha^*(y_\delta)} y\| &\leq \psi(\alpha^*(y_\delta), y_\delta) + f_{\uparrow,y}(\alpha^*(y_\delta)) \\ &\leq \psi(\bar{\alpha}, y_\delta) + f_{\uparrow,y}(\bar{\alpha}). \end{aligned}$$

In the case $\bar{\alpha} \leq \alpha^*(y_\delta)$ we find in a similar manner

$$\|R_{\alpha^*(y_\delta)} y_\delta - R_{\alpha^*(y_\delta)} y\| \leq \|R_{\bar{\alpha}} y_\delta - R_{\bar{\alpha}} y\| \leq \frac{1}{\tilde{C}}\left(\psi(\bar{\alpha}, y_\delta) + f_{\uparrow,y}(\bar{\alpha})\right),$$

and

$$\begin{aligned} \|R_{\alpha^*(y_\delta)} y - A^\dagger y\| &\leq \Phi\left(\psi(\alpha^*(y_\delta), y_\delta) + f_{\downarrow, y-y_\delta}(\alpha^*(y_\delta), y - y_\delta)\right) \\ &\leq \Phi\left(\psi(\bar{\alpha}, y_\delta) + f_{\downarrow, y-y_\delta}(\bar{\alpha}, y - y_\delta)\right). \qquad \square \end{aligned}$$

If upper bounds on $\psi(\alpha, y_\delta)$ similar to (3.11), (3.12) and on $f_{\downarrow, y-y_\delta}(\alpha)$, $f_{\uparrow,y}(\alpha)$ can be found, then convergence rates can be established.

We use the abstract results in this section to study some specific examples of noise level-free parameter choice rules in the next section.

4. Analysis of specific parameter choice rules. In this section we describe some well-known parameter choice rules: since a parameter choice rule is fixed by stating a specific functional ψ, we will in the following always refer to this functional as the parameter choice rule. Some well-known functionals are as follows:

- the *Hanke-Raus rules* with parameter $\tau \in (0, \infty]$,

$$\psi_{HR,\tau}(\alpha, y_\delta) = \sqrt{\frac{1}{\alpha}\int |r_\alpha(\lambda)|^{2+\frac{1}{\tau}} dF_\lambda \|Qy_\delta\|^2}, \tag{4.1}$$

- the *quasi-optimality rule*

$$\psi_{QO}(\alpha, y_\delta) = \sqrt{\int r_\alpha(\lambda)^2 \lambda g_\alpha(\lambda)^2 dF_\lambda \|y_\delta\|^2}, \tag{4.2}$$

- for discrete regularization methods with a fixed sequence $\beta_i < \alpha_i$, $\beta_i \in M_d$, the *discrete quasi-optimality rule*

$$\psi_{DQO}(\alpha_i, y_\delta) = \sqrt{\int (g_{\alpha_i}(\lambda) - g_{\beta_i}(\lambda))^2 \lambda dF_\lambda \|y_\delta\|^2}, \tag{4.3}$$

- the *(modified) L-curve method* with parameter $\mu > 0$,

$$\phi_{\mu L}(\alpha, y_\delta) = \left(\int \lambda g_\alpha(\lambda)^2 dF_\lambda \|y_\delta\|^2\right)^{\frac{\mu}{2}} \sqrt{\int r_\alpha(\lambda)^2 dF_\lambda \|Q y_\delta\|^2}. \tag{4.4}$$

The Hanke-Raus rules were introduced in [11] with a slightly more general definition as here. The parameter τ is fixed and usually (but not necessarily) chosen as the qualification index μ_0 of the method. The special choice $\tau = \infty$ yields a particular simple functional ψ, since if A is injective, $\psi_{HR,\infty} = \frac{1}{\sqrt{\alpha}}\|Ax_{\alpha,\delta} - y_\delta\|$ is nothing but the residual weighted with $\alpha^{-\frac{1}{2}}$.

The quasi-optimality rules ψ_{QO} and ψ_{DQO} are the oldest ones and were introduced by Tikhonov and Glasko [23, 24] for Tikhonov regularization. In these papers, the rules were defined using the functional $\psi(\alpha, y_\delta) = \|\alpha \frac{d}{d\alpha} R_\alpha y_\delta\|$, but for Tikhonov regularization this definition is identical to the one given above. In fact, the form of the functional ψ_{QO} as in (4.2), which does not require the spectral filter to be differentiable, goes back to Neubauer [20]. The rule (4.2) agrees with the classical quasi-optimality rule for Tikhonov regularization, but it should be noted that, e.g., for iterated Tikhonov regularization it is different to $\|\alpha \frac{d}{d\alpha} R_\alpha y_\delta\|$.

The rule ψ_{DQO} can be understood as a discretization of ψ_{QO}. Again this rule goes back to Tikhonov and Glasko [23, 24] who used $\alpha_i = q^i, \beta_i = q^{i+1}$ with $q < 1$ for Tikhonov regularization. With this sequence, it is not difficult to understand ψ_{DQO} as a discrete version of the original quasi-optimality rule of Tikhonov and Glasko, where the derivative is replaced by a difference quotient on a logarithmic scale of α. Moreover, it was shown in [20] that for Landweber iteration, with $\alpha_i = \frac{1}{k}$, $\beta_i = \frac{1}{2k}$, ψ_{DQO} coincides with ψ_{QO}. This justifies the notion of discrete quasi-optimality rule. Further references on quasi-optimality rules can be found in [2, 3].

The L-curve method was introduced by Lawson and Hanson [15] and further studied by Hansen [12] and Hansen and O'Leary [13]. It was cast into the minimization form (4.4) (and generalized to the modified L-curve method) by Regińska [22].

An overview of noise level-free parameter choice rules can be found in [6, 10], and in particular in [21]. Further rules not stated above are listed in [21], e.g., the generalized cross validation [25] and the Brezinksi-Rodrigues-Seatzu rule [5]. A numerical comparison of some rules was preformed in [10], [21], and, recently, in [4]. In [21] also efficient numerical improvements are tested, some of which can be found as well in [8, 9].

We note that the functionals in (4.1)–(4.3) have the form $\psi(\alpha, y_\delta) = \|S_\alpha y_\delta\|$ with an appropriate operator S_α, in particular, they satisfy the subadditivity Condition 3.7.

Let us look at Assumption 2.4 for the above mentioned rules.

PROPOSITION 4.1.

1. *Let the spectral filter g_α be continuous on M with respect to α. Then Assumption 2.4 is satisfied for $\psi_{QO}, \psi_{DQO}, \psi_{\mu L}$.*
2. *If $r_\alpha(\lambda)$ is lower semicontinuous, and if for $\tau \in \mathbb{R}^+ \cup \infty$ a positive number ϵ exists such that*

$$\text{the function } \rho(x) = x^{\frac{1}{2+\frac{1}{\tau}-\epsilon}} \text{ is a qualification.} \tag{4.5}$$

Then Assumption 2.4 is satisfied for $\psi_{HR,\tau}$.

ETNA
Kent State University
http://etna.math.kent.edu

Proof. The conditions (A1)–(A3) are obvious. If $g_\alpha(\lambda)$ is continuous in α then so is $r_\alpha(\lambda)$; from the Lebesgue dominated convergence theorem it follows that $\psi(.,y_\delta)$ is continuous, thus (A4) holds. If r_α is merely lower semicontinuous, (A4) is a consequence of Fatou's lemma for $\psi_{HR,\tau}$. Concerning (A5) we observe that if $z \in D(A^\dagger)$,

$$\psi_{QO}(\alpha,z)^2 = \int r_\alpha(\lambda)^2\lambda^2 g_\alpha(\lambda)^2 dE_\lambda \|A^\dagger z\|^2;$$

and the function $r_\alpha(\lambda)^2\lambda^2 g_\alpha(\lambda)^2$ is uniformly bounded and tends to 0 pointwise, thus by the dominated convergence theorem we obtain that (A5) holds for ψ_{QO}. In a similar way we can show this for $\psi_{\mu L}$. Using (4.5) for $\psi_{HR,\tau}$, by the dominated convergence theorem we obtain

$$\psi_{HR,\tau}(\alpha,z)^2 = \int |r_\alpha(\lambda)|^{2+\frac{1}{\tau}}\frac{\lambda}{\alpha} dE_\lambda \|A^\dagger z\|^2 \leq D_\rho \int |r_\alpha(\lambda)|^{\epsilon} dE_\lambda \|A^\dagger z\|^2,$$

which implies (A5). For ψ_{DQO} we observe that

$$\psi_{QO}(\alpha,z) \leq \|R_{\alpha_i}z - A^\dagger z\|^2 + \|R_{\beta_i}z - A^\dagger z\|^2.$$

Both terms tend to 0 as $\alpha_i \to 0$ $\beta_i \to 0$, thus (A5) holds for ψ_{DQO} as well. □

It is easy to see that all the regularization methods mentioned in Section 2 are continuous except for spectral cutoff. In the continuous case, $M = M_c$, its residual function is not continuous but only lower semicontinuous so that the second case in Proposition 4.1 applies. Of course, since the rule ψ_{QO} is 0 for spectral cutoff, it is of no use here, even though Assumption 2.4 holds (but of course not the Conditions 3.2 and 3.3).

Concerning Condition 3.2 it is obvious that it holds for $\psi_{QO}, \psi_{\mu L}, \psi_{HR,\tau}$ for any regularization for which g_α, r_α are continuous in α and satisfy $g_\alpha(\lambda)r_\alpha(\lambda) \neq 0$ for all λ. If $g_{\beta_i}(\lambda) - g_{\alpha_i}(\lambda) \neq 0$ for all λ, then the same is true for ψ_{DQO}. All this can be shown by Lemma 3.4, which covers already a majority of regularization methods, except Landweber iteration when $\|A\| = 1$ and spectral cutoff. However, these cases are settled by Lemma 3.5.

PROPOSITION 4.2. *For Tikhonov regularization and Landweber iteration, Assumption 2.4 and Condition 3.2 are satisfied for $\psi_{HR,\tau}$ (for all $\tau \in (0,\infty]$), $\psi_{QO}, \psi_{DQO}, \psi_{\mu L}$. For spectral cutoff, Assumption 2.4 and Condition 3.2 are satisfied for $\psi_{HR,\tau}$ (for all $\tau \in (0,\infty]$), and for $\psi_{\mu L}$.*

Proof. We notice that g_α for Tikhonov regularization, Landweber iteration and spectral cutoff satisfies one of the conditions in Proposition 4.1, so that Assumption 2.4 holds in all cases. Moreover, if $A^\dagger y \neq 0$, then for Tikhonov regularization and Landweber iteration with $\|A\| < 1$ for all functionals in the proposition the positivity of ψ, (3.3) holds, thus Lemma 3.4 yields the result in this case. If $A^\dagger y = 0$ or $\|A\| = 1$ for Landweber iteration, and in the case of spectral cutoff with $\psi_{HR,\tau}$, Lemma 3.5 can be applied, where (3.4) can be shown elementary. □

We note that for spectral cutoff, Condition 3.2 is never satisfied for ψ_{QO} and usually (without restrictive condition on $A^\dagger y$) not satisfied for ψ_{DQO}.

4.1. Convergence analysis. We have established all ingredients for the abstract convergence theorem except Condition 3.3. Its verification is at the heart of a convergence proof and turns out to be the most difficult part.

In this section we restrict ourselves to subadditive functionals, i.e., to $\psi_{HR,\tau}$, including the case $\tau = \infty$, and ψ_{QO}, ψ_{DQO}. By the triangle inequality it is obvious that

LEMMA 4.3. *$\psi_{HR,\tau}$ for $\tau \in (0,\infty]$, $\psi_{QO}(\alpha,y_\delta)$, $\psi_{DQO}(\alpha_i,y_\delta)$ satisfy Condition 3.7.*

In view of Lemma 3.11, we can focus on Condition 3.10 and (3.9) to verify Condition 3.3. A condition on the noise such that (3.9) holds was stated in [14] for (iterated) Tikhonov and

ETNA
Kent State University
http://etna.math.kent.edu

the quasi-optimality rule. This has been generalized to other regularizations methods in [20]. We will show that similar conditions are useful for other parameter choice rules as well.

DEFINITION 4.4. *For $p \geq 1$, t_1, $\nu > 0$ fixed, we define the set of restricted noisy data $\mathcal{N}_p$ as*

$$\mathcal{N}_p := \{e \in Y \mid \textit{such that } (4.6) \textit{ holds}\},$$

where

$$t^p \int_t^\infty \lambda^{-1} dF_\lambda \|e\|^2 \leq \nu \int_0^t \lambda^{p-1} dF_\lambda \|e\|^2 \quad \forall 0 < t \leq t_1. \tag{4.6}$$

It is elementary to see that for $p_1 \leq p_2$, the inclusion $\mathcal{N}_{p_2} \subset \mathcal{N}_{p_1}$ is valid. Using the set $\mathcal{N}_p$ we can establish (3.9) for some of the parameter choice rules.

PROPOSITION 4.5. *Let g_α be a spectral filter.*

- *Let (2.9), (2.6) hold and let there exists t_1, ν such that $Q(y_\delta - y) \in \mathcal{N}_1$. Then (3.9) is satisfied for $\psi_{HR,\tau}$ for any $\tau \in (0, \infty]$.*
- *Let (2.9), (2.7) hold and let there exists t_1, ν such that $y_\delta - y \in \mathcal{N}_2$. Then (3.9) is satisfied for ψ_{QO}.*
- *Let (2.9), (2.7) hold and let there exists t_1, ν such that $y - y_\delta \in \mathcal{N}_2$. If additionally a constant $0 < \eta \leq \|A\|$ exist, such that for all $0 < \gamma \leq 1$ there exists a constant $\underline{C}_{DQO,l,\gamma} > 0$ with*

$$|g_{\beta_i}(\lambda) - g_{\alpha_i}(\lambda)| \geq \underline{C}_{DQO,l,\gamma} |g_{\alpha_i}(\lambda)| \quad \forall 0 \leq \lambda \leq \gamma\alpha_i, \quad 0 \leq \gamma\alpha \leq \eta, \tag{4.7}$$

then (3.9) is satisfied for ψ_{DQO}.

Proof. Denote by C a generic constant and fix $\gamma \leq 1$ such that $\gamma\alpha \leq \min\{\eta, t_1\}$, for all $\alpha \in M$. Here, η is always understood as the minimum of the η such that the imposed conditions in (2.6)–(2.9) and (4.7) hold. Then,

$$\begin{aligned}
\psi_{HR,\tau}(\alpha, y - y_\delta)^2 &\geq_{(2.9)} C \frac{1}{\alpha} \int_0^{\gamma\alpha} dF_\lambda \|Q(y - y_\delta)\|^2 \\
&\geq_{(4.6)} C \int_{\gamma\alpha}^\infty \frac{1}{\lambda} dF_\lambda \|Q(y - y_\delta)\|^2 \geq_{(2.2)} C \int_{\gamma\alpha}^\infty \lambda g_\alpha(\lambda)^2 dF_\lambda \|y - y_\delta\|^2.
\end{aligned}$$

Furthermore, we obtain

$$\begin{aligned}
\int_0^{\gamma\alpha} \lambda g_\alpha(\lambda)^2 dF_\lambda \|y - y_\delta\|^2 &\leq_{(2.2)} C \int_0^{\gamma\alpha} |g_\alpha(\lambda)| dF_\lambda \|Q(y - y_\delta)\|^2 \\
&\leq_{(2.6)} C \frac{1}{\alpha} \int_0^{\gamma\alpha} dF_\lambda \|Q(y - y_\delta)\|^2 \leq_{(2.9)} C \frac{1}{\alpha} \int_0^{\gamma\alpha} r_\alpha(\lambda)^{2+\frac{1}{\tau}} dF_\lambda \|Q(y - y_\delta)\|^2.
\end{aligned}$$

Since $\|R_\alpha y - R_\alpha y_\delta\|^2 = \int_0^\infty \lambda g_\alpha(\lambda)^2 dF_\lambda \|y - y_\delta\|^2$ we have shown the result for the $\psi_{HR,\tau}$. For ψ_{QO} we observe that

$$\begin{aligned}
\psi_{QO}^2(\alpha, y - y_\delta) &\geq_{(2.9)} C \int_0^{\gamma\alpha} \lambda g_\alpha(\lambda)^2 dF_\lambda \|y - y_\delta\|^2 \geq_{(2.7)} C \frac{1}{\alpha^2} \int_0^{\gamma\alpha} \lambda dF_\lambda \|y - y_\delta\|^2 \\
&\geq_{(4.6)} C \int_{\gamma\alpha}^\infty \frac{1}{\lambda} dF_\lambda \|y - y_\delta\|^2 \geq_{(2.2)} C \int_{\gamma\alpha}^\infty \lambda g_\alpha(\lambda)^2 dF_\lambda \|y - y_\delta\|^2.
\end{aligned}$$

On the other hand, we get

$$\int_0^{\gamma\alpha} \lambda g_\alpha(\lambda)^2 dF_\lambda \|y - y_\delta\|^2 \leq_{(2.9)} C \int_0^{\gamma\alpha} \lambda g_\alpha(\lambda)^2 r_\alpha(\lambda)^2 dF_\lambda \|y - y_\delta\|^2,$$

ETNA
Kent State University
http://etna.math.kent.edu

which shows the result for ψ_{QO}. Finally, for ψ_{DQO} we can estimate

$$\begin{aligned}\psi_{DQO}^2(\alpha, y-y_\delta) &\geq_{(4.7)} C\int_0^{\gamma\alpha_i}\lambda(g_{\alpha_i}(\lambda))^2 dF_\lambda\|y-y_\delta\|^2\\ &\geq_{(2.7)} C\frac{1}{\alpha_i^2}\int_0^{\gamma\alpha_i}\lambda dF_\lambda\|y-y_\delta\|^2 \geq_{(4.6)} C\int_{\gamma\alpha_i}^\infty \frac{1}{\lambda}dF_\lambda\|y-y_\delta\|^2\\ &\geq_{(2.2)} C\int_{\gamma\alpha}^\infty \lambda(g_{\alpha_i}(\lambda))^2 dF_\lambda\|y-y_\delta\|^2.\end{aligned}$$

As before, we get

$$\int_0^{\gamma\alpha_i}\lambda(g_{\alpha_i}(\lambda))^2 dF_\lambda\|y-y_\delta\|^2 \leq_{(4.7)} C\int_0^{\gamma\alpha}\lambda(g_{\beta_i}(\lambda)-g_{\alpha_i}(\lambda))^2 dF_\lambda\|y-y_\delta\|^2. \qquad \square$$

This result implies the convergence of the mentioned parameter choice rules. Concerning (4.7) it will be shown below that if there is a $q < 1$ such that

$$\beta_i \leq q\alpha_i \quad \forall i, \tag{4.8}$$

then (4.7) holds for the (discrete) Tikhonov regularization and for Landweber iteration. This gives the convergence result for the example regularization methods in this paper.

THEOREM 4.6. *Let $y_\delta - y \in \mathcal{N}_2$, and for ψ_{DQO} let (4.8) hold. Then the parameter choices $\psi_{QO}, \psi_{DQO}, \psi_{HR,\tau}$ converge for Tikhonov regularization and Landweber iteration. Let $y_\delta - y \in \mathcal{N}_1$, then the parameter choice with $\psi_{HR,\tau}$ converges for Tikhonov regularization, Landweber iteration, and spectral cutoff.*

Proof. The conditions (2.9), (2.6) and (2.7) hold for Tikhonov regularization and Landweber iteration. For spectral cutoff, (2.6) and (2.9) hold. Altogether, convergence follows from Propositions 4.2, 4.5 and Theorem 3.6. It remains to prove that (4.7) is implied by (4.8) for Tikhonov regularization and Landweber iteration. For Tikhonov regularization, inequality (4.7) holds for $\lambda \in [0, \gamma\alpha_i]$ if $\alpha_i - \beta_i \geq \underline{C}_{DQO,l,\gamma}(\gamma\alpha_i + \beta_i)$ is satisfied. Condition (4.8) suffices for this inequality with $\underline{C}_{DQO,l,\gamma} = \frac{\gamma+q}{1-q}$. For Landweber iteration, the monotonicity in λ of the left hand side in (4.7) yields for $\lambda \in [0, \gamma\alpha_i]$

$$\begin{aligned}g_{\frac{1}{\beta_i}}(\lambda) - g_{\frac{1}{\alpha_i}}(\lambda) \geq g_{\frac{1}{\beta_i}}(\gamma\alpha_i) - g_{\frac{1}{\alpha_i}}(\gamma\alpha_i) &= \frac{1}{\gamma\alpha_i}\left((1-\gamma\alpha_i)^{\frac{1}{\alpha_i}} - (1-\gamma\alpha_i)^{\frac{1}{\beta_i}}\right)\\ &\geq \frac{1}{\gamma\alpha_i}\left((1-\gamma\alpha_i)^{\frac{1}{\alpha_i}} - (1-\gamma\alpha_i)^{\frac{1}{q\alpha_i}}\right)\\ &= \frac{1}{\gamma\alpha_i}\left(\left((1-\gamma\alpha_i)^{\frac{1}{\gamma\alpha_i}}\right)^\gamma - \left((1-\gamma\alpha_i)^{\frac{1}{\gamma\alpha_i}}\right)^{\frac{\gamma}{q}}\right).\end{aligned}$$

Take $\eta = \frac{1}{2}$, so that $0 \leq \gamma\alpha_i$ implies $(1-\gamma\alpha_i)^{\frac{1}{\gamma\alpha_i}} \in [\frac{1}{4}, \frac{1}{e}]$, which yields the estimate

$$g_{\frac{1}{\beta_i}}(\lambda) - g_{\frac{1}{\alpha_i}}(\lambda) \geq \frac{1}{q}\left(\min_{x\in[\frac{1}{4},\frac{1}{e}]}(x^\gamma - x^{\frac{\gamma}{q}})\right)\frac{1}{\alpha_i} \geq C\frac{1}{\alpha_i},$$

with a positive constant C. With (2.6) we arrive at (4.7). $\square$

For the quasi-optimality rule, this theorem and the previous propositions have been shown in [20] and for (iterated) Tikhonov regularization in [14]. The new results in this paper are extensions to $\psi_{HR,\tau}, \psi_{DQO}$.

4.2. Convergence rate analysis. Let us now consider the conditions for convergence rates and optimal order convergence using Theorem 3.12. Since in this section we only consider ψ_{QO}, $\psi_{HR,\tau}$, ψ_{DQO}, we can use subadditivity (Condition 3.7). In the previous section, we have already considered premises when (3.9) of Condition 3.10 is satisfied. This section is concerned with Condition 3.8, in particular (3.7), and the upper estimates (3.11) and (3.12). For the optimal order theorem we additionally have to show (3.14), (3.15). More precisely, we establish conditions for the following estimates for the functionals ψ with some generic constants C and index function Φ:

$$\psi(\alpha, y_\delta - y) \leq C\frac{\delta}{\sqrt{\alpha}} \quad \forall \alpha \in M, \tag{4.9}$$

$$\psi(\alpha, y) \leq C\|R_\alpha y - A^\dagger y\| \quad \forall \alpha \in M, \tag{4.10}$$

$$\psi(\alpha, y_\delta - y) \leq C\|R_\alpha y_\delta - R_\alpha y\| \quad \forall \alpha \in M, \tag{4.11}$$

$$\psi(\alpha, y) \geq \Phi^{-1}\|R_\alpha y - A^\dagger y\| \quad \forall \alpha \in M. \tag{4.12}$$

For some of these inequalities we need some additional conditions. We list them for later reference. For the discrete quasi-optimality rule we require the following in addition to (4.7)

$$|g_{\beta_i}(\lambda) - g_{\alpha_i}(\lambda)| \leq \overline{C}_{DQO,c}|g_{\alpha_i}(\lambda)| \quad \forall \lambda \in (0, \|A\|^2], \tag{4.13}$$

$$|r_{\beta_i}(\lambda) - r_{\alpha_i}(\lambda)| \leq \overline{D}_{DQO,c}|r_{\alpha_i}(\lambda)| \quad \forall \lambda \in (0, \|A\|^2], \tag{4.14}$$

and that there exists a constant $0 < \eta < \|A\|^2$ such that for all $0 < \gamma \leq 1$

$$|r_{\beta_i}(\lambda) - r_{\alpha_i}(\lambda)| \geq \underline{D}_{DQO,h}|r_{\alpha_i}(\lambda)| \quad \forall \|A\|^2 \geq \lambda \geq \gamma\alpha_i \quad 0 \leq \gamma\alpha_i \leq \eta. \tag{4.15}$$

We furthermore need conditions on the exact solution: $x^\dagger = A^\dagger y$:

- there exists a constant $\eta > 0$ and an index function such that for all $0 \leq t \leq \eta$,

$$\int_0^t r_t(\lambda)^2 dE_\lambda \|x^\dagger\|^2 \leq \Psi\left(\int_t^{\|A\|^2} r_t(\lambda)^2 dE_\lambda \|x^\dagger\|^2\right), \tag{4.16}$$

- there exists a constant $\eta > 0$ and an index function such that for all $0 \leq t \leq \eta$,

$$\int_0^t r_t(\lambda)^2 dE_\lambda \|x^\dagger\|^2 \leq \Psi\left(\frac{1}{t}\int_t^{\|A\|^2} r_t(\lambda)^2 \lambda dE_\lambda \|x^\dagger\|^2\right), \tag{4.17}$$

- there exists a constant θ_1 such that for all $t \in M$,

$$\int_t^\infty r_t^2(\lambda)\lambda d\|x^\dagger\|^2 \leq \theta_1\, t\int_t^\infty r_t^2(\lambda) d\|x^\dagger\|^2. \tag{4.18}$$

Note that (4.16) implies (4.17) with the same Ψ. However, using (4.17) enables us in the case of $\psi_{HR,\infty}$ to get a better estimate than that for (4.16).

The Hanke-Raus rules require an additional condition on the noise besides (4.6): There exists a constant η and θ_2 such that for all $0 \leq t \leq \eta$,

$$\int_0^t dF_\lambda \|Qy_\delta - y\|^2 \leq \theta_2\frac{1}{t}\int_0^t \lambda dF_\lambda \|Qy_\delta - y\|^2. \tag{4.19}$$

The first lemma concerns (4.9).

LEMMA 4.7. *Let g_α be a spectral filter. Then (4.9) holds*

- *for $\psi_{HR,\tau}$ with any $\tau \in (0,\infty]$,*
- *for ψ_{QO} if (2.6) is satisfied,*
- *for ψ_{DQO} if (2.6) and (4.13) are satisfied.*

Proof. This follows from $r_\alpha(\lambda) \leq (1 + C_g)$, $|\lambda g_\alpha(\lambda)| \leq C_g$, and (2.6). In the case of ψ_{DQO} we additionally need (4.13). ∎

Now we proceed to show the estimate (4.10).

LEMMA 4.8. *Let g_α be a spectral filter. Then (4.10) holds*

- *for $\psi_{HR,\tau}$ if $\tau < \infty$ and $\rho(x) = x^\tau$ is a qualification,*
- *for $\psi_{HR,\infty}$, if (4.18) holds,*
- *for ψ_{QO},*
- *for ψ_{DQO} if (4.14) is satisfied.*

Proof. For ψ_{QO}, the result follows from (2.2) and for $\psi_{HR,\tau}$ from the qualification condition $r_\alpha(\lambda)^{\frac{1}{\tau}}|\frac{\lambda}{\alpha}| \leq C_\rho$. For ψ_{DQO}, it follows from

$$\lambda^2|g_{\alpha_i}(\lambda) - g_{\beta_i}(\lambda)|^2 = |r_{\alpha_i}(\lambda) - r_{\beta_i}(\lambda)|^2,$$

together with (4.14). For $\psi_{HR,\infty}$, we obtain by (4.18)

$$\begin{aligned}\psi_{HR,\infty}(\alpha, Ax^\dagger)^2 &\leq \int_0^\alpha r_\alpha^2(\lambda)dE_\lambda\|x^\dagger\|^2 + \int_\alpha^{\|A\|^2} r_\alpha^2(\lambda)\frac{\lambda}{\alpha}dE_\lambda\|x^\dagger\|^2\\ &\leq_{(4.18)} \int_0^\alpha r_\alpha^2(\lambda)dE_\lambda\|x^\dagger\|^2 + \theta_1\int_\alpha^{\|A\|^2} r_\alpha^2(\lambda)dE_\lambda\|x^\dagger\|^2\\ &\leq \max\{1,\theta_1\}\|R_\alpha y - A^\dagger x^\dagger\|^2. \quad ∎\end{aligned}$$

Concerning (4.11) we have the following result.

LEMMA 4.9. *Let g_α be a spectral filter. Then (4.11) holds*

- *for $\psi_{HR,\tau}$ with $\tau \in (0,\infty]$, if $\rho(x) = x^{\frac{\tau}{2\tau+1}}$ is a qualification, (2.7) holds, (4.19) is satisfied, and $y_\delta - y \in \mathcal{N}_1$,*
- *for ψ_{QO},*
- *for ψ_{DQO}, if (4.13) holds.*

Proof. The result for ψ_{QO} follows immediately from $|r_\alpha(\lambda)| \leq 1 + C_g$ and for ψ_{DQO} from (4.13). For the Hanke-Raus rules, we choose γ such that $\gamma\alpha \leq \min\{t_1,\eta\}$, where η is the minimum of the constants in (4.19), (2.7). Then

$$\begin{aligned}&\psi_{HR,\tau}(\alpha, y_\delta - y)\\ &\quad\leq_{(2.2),(2.10)} (1 + C_g)^{2+\frac{1}{\tau}}\frac{1}{\alpha}\int_0^{\gamma\alpha} dF_\lambda\|Qy_\delta - y\|^2 + C_{3,\rho}\int_{\gamma\alpha}^\infty \frac{1}{\lambda}dF_\lambda\|Qy_\delta - y\|^2\\ &\quad\leq_{(4.6)} C\frac{1}{\alpha}\int_0^{\gamma\alpha} dF_\lambda\|Qy_\delta - y\|^2 \leq_{(4.19)} C\frac{1}{\alpha^2}\int_0^{\gamma\alpha}\lambda dF_\lambda\|Qy_\delta - y\|^2\\ &\quad\leq_{(2.7)} C\int_0^{\gamma\alpha}\lambda g_\alpha(\lambda)^2 dF_\lambda\|Qy_\delta - y\|^2 \leq \|R_\alpha(y_\delta - y)\|^2. \quad ∎\end{aligned}$$

Finally we come to (4.12).

LEMMA 4.10. *Let g_α be a spectral filter. Then (4.12) holds*

- *for $\psi_{HR,\tau}$, if $\tau < \infty$, $\tau = \mu_0$, where μ_0 is the qualification index, and (4.16) is satisfied,*
- *for $\psi_{HR,\infty}$, if $\tau = \infty$ and (4.17) is satisfied,*
- *for ψ_{QO}, if (2.8) and (4.16) is satisfied,*

ETNA
Kent State University
http://etna.math.kent.edu

- *for ψ_{DQO}, if (4.16) and (4.15) holds.*

The index function Φ in (4.12) is in all cases related to the index functions (4.16) or (4.17), respectively, by

$$\Phi(x) = \sqrt{\Psi(C_1 x^2) + C_2 x^2},$$

where C_1, C_2 are constants.

Proof. Let $\gamma \leq 1$ be such that $\gamma\alpha \leq \eta$ with η being the minimum of the constants appearing in the required conditions. For $\psi_{HR,\tau}$ with τ being the qualification index, and for ψ_{QO}, ψ_{DQO}, we obtain lower bounds

$$\psi_{HR,\mu_0}(\alpha, Ax^\dagger)^2 \geq_{(2.11)} D_{ind,\mu_0} \int_{\gamma\alpha}^{\|A\|^2} r_\alpha(\lambda)^2 dF_\lambda \|x^\dagger\|^2,$$

$$\psi_{QO}(\alpha, Ax^\dagger)^2 \geq_{(2.8)} \underline{C}_{h,\gamma}\gamma^2 \int_{\gamma\alpha}^{\|A\|^2} r_\alpha(\lambda)^2 dF_\lambda \|x^\dagger\|^2,$$

$$\psi_{DQO}(\alpha, Ax^\dagger)^2 \geq \int_{\gamma\alpha_i}^{\|A\|^2} \lambda^2 (g_{\beta_i}(\lambda) - g_{\alpha_i}(\lambda))^2 dE_\lambda \|x^\dagger\|^2$$

$$= \int_{\gamma\alpha_i}^{\|A\|^2} |r_{\beta_i}(\lambda) - r_{\alpha_i}(\lambda)|^2 dE_\lambda \|x^\dagger\|^2 \geq \underline{D}_{DQO,h}^2 \int_{\gamma\alpha_i}^{\|A\|^2} r_{\alpha_i}^2(\lambda) dE_\lambda \|x^\dagger\|^2.$$

On the other hand, from (4.16) with $\gamma \leq 1$, we get

$$\int r_\alpha(\lambda)^2 dE_\lambda \|x^\dagger\|^2 \leq \int_0^\alpha r_\alpha(\lambda)^2 dE_\lambda \|x^\dagger\|^2 + \int_\alpha^{\|A\|^2} r_\alpha(\lambda)^2 dE_\lambda \|x^\dagger\|^2$$

$$\leq_{(4.16)} (\Psi + id)\left(\int_\alpha^{\|A\|^2} r_\alpha(\lambda)^2 dE_\lambda \|x^\dagger\|^2\right)$$

$$\leq (\Psi + id)\left(\int_{\gamma\alpha}^{\|A\|^2} r_\alpha(\lambda)^2 dE_\lambda \|x^\dagger\|^2\right),$$

which establishes the result for three cases. For $\psi_{HR,\infty}$, we can estimate

$$\psi_{HR,\infty}(\alpha, Ax^\dagger)^2 \geq \frac{1}{\alpha} \int_\alpha^{\|A\|^2} r_\alpha(\lambda)^2 \lambda dF_\lambda \|x^\dagger\|^2,$$

and with (4.17) in place of (4.16), and $\lambda \geq \alpha$, the result follows also in this case. ☐

We note that the estimates in these lemmas have been proven by Neubauer [20] for the case of ψ_{QO}.

5. Case studies. In this section we discuss the convergence rate result for the typical cases of a) a regularization method with finite qualification index; b) the case of Landweber iteration, which does not have a finite qualification index, but still has a generalized saturation of qualification; and c) the case of spectral cutoff, which does not show such a saturation. Since convergence rates are impossible without a source condition, in the following we will impose a Hölder type source condition

$$A^\dagger y \in R((A^*A)^\mu). \tag{5.1}$$

ETNA
Kent State University
http://etna.math.kent.edu

5.1. Methods with finite qualification index. We consider the case that $\mu_0 < \infty$ is a qualification index of the regularization method. Tikhonov regularization is a typical example of the methods we discuss here. The convergence of these methods is a simple consequence of Theorem 3.6 and we omit the details here. Let us just mention that Condition (4.5) for $\psi_{HR,\tau}$, which was sufficient for Assumption 2.4, is only a restriction on τ for a low saturating method with $\mu_0 < \frac{1}{2}$, and τ can be chosen arbitrarily otherwise, in particular, the latter is the case for Tikhonov regularization.

The main benefit of a finite qualification index comes within the convergence rate analysis, because (4.16) and (4.17) is satisfied for all $x^\dagger$ with a function Ψ depending on the qualification index. We have the following theorem, extending the results of [20].

THEOREM 5.1. *Let g_α be a continuous, monotone spectral filter, with finite qualification index μ_0 such that $\rho(\lambda) = \lambda^\mu$ is a qualification for all $\mu \in (0, \mu_0]$. Moreover, assume that $x^\dagger$ satisfies a source condition with some $0 < \mu \leq \mu_0$ and let $\int_\eta^{\|A\|^2} dE_\lambda \|x^\dagger\|^2 \neq 0$ for some $\eta > 0$. The following convergence rate estimate is valid*

$$\|R_{\alpha^*(y_\delta)} y_\delta - x^\dagger\| \leq C\delta^{\frac{2\mu}{2\mu+1}\frac{\mu}{\mu_0}} \tag{5.2}$$

in the following cases,
- *for $\psi_{HR,\tau}$, if (2.9), (2.6) hold, $y_\delta - y \in \mathcal{N}_1$, and with the choice $\tau = \mu_0$,*
- *for ψ_{QO}, if (2.9), (2.6), (2.7), (2.8) hold, and $y_\delta - y \in \mathcal{N}_2$,*
- *for ψ_{DQO}, if (2.9), (2.6), (2.7) hold, $y_\delta - y \in \mathcal{N}_2$, and if (4.7), (4.13), (4.14), (4.15) holds.*

Moreover, let (2.9), (2.6) hold and $y_\delta - y \in \mathcal{N}_1$. If $\mu_0 > \frac{1}{2}$, then we obtain for $\psi_{HR,\infty}$ with $\tilde{\mu} = \min\{\mu, \mu_0 - \frac{1}{2}\}$

$$\|R_{\alpha^*(y_\delta)} y_\delta - x^\dagger\| \leq C\delta^{\frac{2\tilde{\mu}}{2\tilde{\mu}+1}\min\{\frac{\tilde{\mu}-\frac{1}{2}}{\mu_0-\frac{1}{2}},1\}}.$$

Proof. We apply Theorem 3.12. According to Proposition 4.1, Assumption 2.4 is automatically satisfied for ψ_{QO}, ψ_{DQO}, ψ_{HR,μ_0} and for $\psi_{HR,\infty}$ by $\mu_0 > \frac{1}{2}$. Condition 3.7 trivially holds. Condition 3.10 was shown in Proposition 4.5. Lemma 4.7 implies that (3.12) holds in the respective cases with $\rho_{\downarrow,y_\delta - y}(\alpha) = \frac{\delta}{\sqrt{\alpha}}$. The source condition (5.1) implies the estimate

$$\|R_\alpha - y_\delta\| \leq C\alpha^\mu \tag{5.3}$$

for $\mu \leq \mu_0$. Together with Lemma 4.8 we obtain (3.11) with $\rho_{\uparrow,y}(\alpha) = \alpha^\mu$ for all cases except $\psi_{HR,\infty}$. However, with (5.1) it is standard that for $\psi_{HR,\infty}$ $\rho_{\uparrow,y}(\alpha) = \alpha^\mu$ for $\mu \leq \mu_0 - \frac{1}{2}$.

It remains to verify Condition 3.8. This is a consequence of Lemma 4.10. We only have to show that (4.16), or (4.17) (in the case $\psi_{HR,\infty}$) holds. However, this is a consequence of (5.3) and the following argument: With the qualification assumption, (4.16) is satisfied since

$$\int_0^{\gamma\alpha} r_\alpha(\lambda)^2 dE_\lambda \|x^\dagger\|^2 \leq C\alpha^{2\mu},$$

while

$$\int_{\gamma\alpha}^{\|A\|^2} r_\alpha(\lambda)^2 dE_\lambda \|x^\dagger\|^2 \geq \alpha^{2\mu_0} \frac{1}{\eta^{2\mu_0}} \int_\eta^{\|A\|^2} dE_\lambda \|x^\dagger\|^2;$$

thus, (4.16) holds with $\Psi(x) = Cx^{\frac{\mu}{\mu_0}}$. Moreover, a similar argument shows that (4.17) holds with $\Psi(x) = Cx^{\frac{\mu}{\mu_0 - \frac{1}{2}}}$. Using Lemma 4.10, we obtain for ψ_{QO}, ψ_{HR,μ_0}, ψ_{DQO} that

ETNA
Kent State University
http://etna.math.kent.edu

Condition 3.8 holds with $\Psi = Cx^{\min\{\frac{\mu}{\mu_0},1\}}$ and for $\psi_{HR,\infty}$ with $\Psi(x) = Cx^{\min\{\frac{\mu-\frac{1}{2}}{\mu_0-\frac{1}{2}},1\}}$. The infimum in Theorem 3.12 can easily be calculated by balancing the terms, which finally yields the result. $\square$

This theorem shows that in general we only get suboptimal rates for heuristic parameter choice rules. Optimal order rates are established if the index of the source condition equals the qualification index μ_0, or, for the $\psi_{HR,\infty}$ rule, equals $\mu_0 - \frac{1}{2}$. We can therefore interpret (4.16) or (4.17) as a quantitative condition on how far the smoothness of $x^\dagger = A^\dagger y$ is away from the qualification index of the method. Note that while the quasi-optimality rules and the Hanke-Raus rules with $\tau = \mu_0$ behave similarly, the Hanke Raus rule with $\tau = \infty$ behaves differently. Depending on the smoothness it is possible that the latter rule might gives better result than the others, while in other cases this situation might be reversed.

A similar analysis is of course possible for other types of source conditions, like logarithmic ones. Since in this case the solution smoothness is much weaker than the qualification index, we expect very weak convergence rates if no further condition on $x^\dagger$ are used.

Let us now turn to optimal order convergence. In the case of finite qualification, the condition (4.16) with $\Psi(x) = x$ reduces to

$$\int_0^t \frac{1}{\lambda^{2\mu}} dE_\lambda \|x^\dagger\|^2 \leq C t^{2(\mu_0-\mu)} \int_t^\infty \frac{1}{\lambda^{2\mu_0}} dE_\lambda \|x^\dagger\|^2, \quad \forall 0 < t < \eta, \tag{5.4}$$

and (4.17) with $\Psi(x) = x$ reduces to

$$\int_0^t \frac{1}{\lambda^{2\mu}} dE_\lambda \|x^\dagger\|^2 \leq C t^{2(\mu_0-\mu)-1} \int_t^\infty \frac{1}{\lambda^{2\mu_0-1}} dE_\lambda \|x^\dagger\|^2, \quad \forall 0 < t < \eta. \tag{5.5}$$

THEOREM 5.2. *Let g_α be a continuous monotone spectral filter, with finite qualification index μ_0 such that $\rho(\lambda) = \lambda^\mu$ is a qualification for all $\mu \in (0, \mu_0]$.*

Assume that $x^\dagger$ satisfies a source condition (5.1) and the decay condition (5.4). Then the optimal order estimate

$$\|R_{\alpha^*(y_\delta)} y_\delta - x^\dagger\| \leq C \inf_\alpha \|R_{\alpha^*(y_\delta)} y_\delta - R_{\alpha^*(y_\delta)} y\| + \|R_{\alpha^*(y_\delta)} y - x^\dagger\| \tag{5.6}$$

holds in the following cases,

- *for ψ_{QO}, if (2.9), (2.7), (2.8) hold, and $y_\delta - y \in \mathcal{N}_2$,*
- *for ψ_{DQO}, if (2.9), (2.7), hold, $y_\delta - y \in \mathcal{N}_2$, and additionally (4.7), (4.13), (4.14), (4.15), hold,*
- *for $\psi_{HR,\tau}$, if (2.9), (2.6) hold, $Q(y_\delta - y) \in \mathcal{N}_1$, $\tau = \mu_0$, and (4.19) holds.*

Moreover, if $x^\dagger$ satisfies a source condition and (5.5), $\mu_0 \geq \frac{1}{2}$, (2.9), (2.6) (4.19), and (4.18) hold, then $\psi_{HR,\infty}$ satisfies the optimal order estimate (5.6).

Proof. This theorem is a consequence of Theorem 3.13, Propositions 4.1 and 4.5, and Lemmas 4.8, 4.9, and 4.10. $\square$

We notice that ψ_{QO} and ψ_{DQO} need similar conditions. For ψ_{HR,μ_0} we need an additional noise condition (4.19), which means that the noise should not be too irregular. This condition is not needed for the quasi-optimality rules.

In the case of $\psi_{HR,\infty}$ we need the rather restrictive condition (4.18). This condition can only be satisfied if the smoothness index of the source condition obeys the inequality $\mu \leq \mu_0 - \frac{1}{2}$.

We also notice that if for some constants $p, q < 1$,

$$p\alpha_i \leq \beta_i \leq q\alpha_i \quad \forall i$$

is valid, then all the conditions (4.7), (4.13), (4.14), (4.15) for ψ_{DQO} are satisfied in the case of Tikhonov regularization.

ETNA
Kent State University
http://etna.math.kent.edu

5.2. Landweber iteration. We now discuss the convergence rate results for Landweber iteration.

The results on convergence rates are different from the previous sections. Note that Landweber iteration does not have a finite qualification index; every function x^μ is a qualification. For Landweber iteration, however, we can find a substitute for the qualification index. Indeed, it can be shown that if $\|A\| < 1$ then

$$|r_\alpha(\lambda)| \geq e^{-C\frac{1}{\alpha}}, \quad 1 > \|A\|^2 > \lambda \geq \gamma\alpha, \tag{5.7}$$

for some constant C. The convergence rates for Landweber iteration with a source condition (5.1) are established in the following theorem.

THEOREM 5.3. *Let $\|A\| < 1$ and the regularization method be Landweber iteration and let $x^\dagger$ satisfy a source condition (5.1). Then the following convergence rate holds with constants C_1, C_2,*

$$\|R_{\alpha^*(y_\delta)} y_\delta - x^\dagger\| \leq C_1 \left(\frac{C_2}{-\log(\delta)} \right)^\mu$$

in the following cases,

- *for $\psi_{HR,\tau}$, if $Q(y_\delta - y) \in \mathcal{N}_1$, for any $\tau \in (0, \infty]$,*
- *for ψ_{QO}, if $y_\delta - y \in \mathcal{N}_2$,*
- *for ψ_{DQO}, if $y_\delta - y \in \mathcal{N}_2$, and constants $p, q < 1$ exists with*

$$p\alpha_i \leq \beta_i \leq q\alpha_i \quad \forall i.$$

Proof. Let us first assume that (4.7), (4.13), (4.14) and (4.15) hold for ψ_{DQO}. As in the previous section we can establish (4.9), and with a source condition it follows easily that $\psi(\alpha, y) \leq C\alpha^\mu$ in all cases (for ψ_{DQO} we need (4.14) here). Proposition 4.5 implies (3.9) under the stated noise conditions. In view of Theorem 3.12 it remains to show (3.7) for $w = Ax^\dagger$. From the source condition it follows that $\|R_\alpha y - A^\dagger y\| \leq C\alpha^\mu$. Using that (2.8) holds for Landweber iteration, (or (4.15) for ψ_{DQO}) it can be shown that in all cases the inequality

$$\psi(\alpha, y)^2 \geq \int_{\gamma\alpha} r_\alpha(\lambda)^\zeta dE_\lambda \|x^\dagger\|^2$$

is valid with $\zeta \in \{2, 2 + \frac{1}{\tau}\}$, depending on the method. Now using (5.7) yields that

$$\psi(\alpha, y) \geq C_3 e^{-\frac{C_4}{\alpha}}.$$

We thus can chose $\Phi^{-1}(x) = C_3 \exp(-\frac{C_4}{x^{1/\mu}})$ to get (3.7). Let us now verify all the required conditions for ψ_{DQO}. In Theorem 4.6 we already showed that (4.8) implies (4.7). From

$$|r_{\beta_i}(\lambda) - r_{\alpha_i}(\lambda)| = r_{\alpha_i}(\lambda)(1 - (1-\lambda)^{\frac{1}{\beta_i} - \frac{1}{\alpha_i}}),$$

(4.14) follow straightforwardly from $\beta_i < \alpha_i$. We observe that for $1 > \|A\|^2 < \lambda \geq \gamma\alpha$,

$$1 - (1-\lambda)^{\frac{1}{\beta_i} - \frac{1}{\alpha_i}} \geq 1 - (1-\gamma\alpha_i)^{(\frac{1}{q}-1)\frac{1}{\alpha_i}} \geq 1 - \left((1-\gamma\alpha_i)^{\frac{1}{\gamma\alpha_i}} \right)^{\gamma(\frac{1}{q}-1)},$$

and (4.15) follows as in the proof of Theorem 4.6. The inequality (4.13) is equivalent to

$$\frac{(1-\lambda)^{\frac{1}{\alpha_i}} - (1-\lambda)^{\frac{1}{\beta_i}}}{1 - (1-\lambda)^{\frac{1}{\alpha_i}}} \leq \overline{C}_{DQO,c},$$

and with $\beta_i \geq p\alpha_i$ the left-hand side can be bounded by

$$\frac{(1-\lambda)^{\frac{1}{\alpha_i}} - (1-\lambda)^{\frac{1}{\beta_i}}}{1-(1-\lambda)^{\frac{1}{\alpha_i}}} \leq \frac{(1-\lambda)^{\frac{1}{\alpha_i}} - (1-\lambda)^{\frac{1}{q\alpha_i}}}{1-(1-\lambda)^{\frac{1}{\alpha_i}}}.$$

Since $(1-\lambda)^{\frac{1}{\alpha_i}} \in (0,1]$ and $\frac{x-x^{\frac{1}{q}}}{1-x}$ with $q < 1$ is uniformly bounded for $x \in [0,1]$, (4.13) is proven. ◻

This shows that for methods with high or infinite qualification index, only slow convergence rates can be expected. Of course, with the appropriate conditions, we can prove again optimal order rates in the same line as before. However, in the case of Landweber iteration, the decay conditions (4.16) are more restrictive compared to the finite qualification case.

5.3. Spectral cutoff. The case of the spectral cutoff (or truncated singular value decomposition) is different in several aspects. The quasi-optimality rule ψ_{QO} is not applicable here, because it is 0 (Condition 3.2 is violated). Also, the discrete version ψ_{DQO} is not appropriate, because Condition 3.2 cannot be verified in general. However, Hanke-Raus rules can be analyzed. We notice that in the case of spectral cutoff, $\psi_{HR,\tau}$ is independent of τ.

The convergence rate result are stated in the following theorem.

THEOREM 5.4. *Let $Q(y_\delta - y) \in \mathcal{N}_1$ and the regularization method be spectral cutoff.*

- *If a source condition is satisfied and if there is an index function such that*

$$\int_0^t dE_\lambda \|x^\dagger\|^2 \leq \Phi\left(\frac{1}{t}\int_0^t \lambda dE_\lambda \|x^\dagger\|^2\right) \qquad \forall t \in (0,\alpha_0], \tag{5.8}$$

  then for $\psi_{HR,\tau}$ we obtain

$$\|R_{\alpha^*(y_\delta)} y_\delta - x^\dagger\| \leq \Phi(\delta^{\frac{2\mu}{1+2\mu}}).$$

- *If (5.8) is satisfied with $\Psi(x) = x$, then the optimal error bound (5.6) holds.*

Proof. For spectral cutoff, (2.6) and (2.9) holds so that by Corollary 4.2, Proposition 4.5, Lemmas 4.7, 4.8, and 4.9 all conditions of Theorem 3.12 and Theorem 3.13 are satisfied except for (3.7). However, the condition (5.8), which is used instead of (4.16) or (4.17) is exactly (3.7). ◻

The difference to the previous methods is that for spectral cutoff we cannot proof convergence rates only by using a source condition. The reason is, as already mentioned, the lack of a saturation of qualification.

5.4. The L-curve method. Let us finally discuss the drawback of the L-curve method. We want to argue that the analysis in this paper does not apply to this method. We have seen that the central ingredient for a convergence proof is a noise condition (Condition 3.3). For convergence rates a quantitative version has to be used. Since the L-curve is not subadditive (Condition 3.7 does not hold) condition (3.16) was used to prove convergence or convergence rates. We now argue that such an estimate does not hold in realistic cases. First, let us note that one has some freedom in choosing ψ: if ψ satisfies Assumption 2.4, then so does $\Phi(\psi)$, where Φ is an index function with $\Phi(\infty) = \infty$. The choice of Φ, however, should be related to estimates such as (3.16) in the sense that the left- and right-hand sides should have the same degree of homogeneity. The left-hand side in (3.16) is homogeneous of degree 1 in the noise $y_\delta - y$: $\|R_\alpha \lambda(y_\delta - y)\| = \lambda \|R_\alpha(y_\delta - y)\|$, which suggests that ψ should be homogeneous of degree 1 as well. If this condition holds, then the set $\mathcal{N}_z$ is scaling invariant, which means that the noise condition $y_\delta - y \in \mathcal{N}_y$ does not need any information on the noise level $\|y_\delta - y\|$.

ETNA
Kent State University
http://etna.math.kent.edu

With this argument it is clear that the functional for the (modified) L-curve method should be homogeneous of degree 1, i.e.,

$$\psi_{\mu L,scaled}(\alpha, y_\delta) = \psi_{\mu L}(\alpha, y_\delta)^{\frac{1}{1+\mu}}, \quad \mu > 0.$$

However, for the scaled version the bound (3.16) does not hold.

PROPOSITION 5.5. *Let g_α be a spectral filter. For the functional $\psi_{\mu L,scaled}$ there cannot be a set $\mathcal{N}_z$ with $\mathcal{N}_z \cap D(A^\dagger) = \emptyset$ such that (3.16) holds for all $\alpha \in M$ with $C > 0$ and $f(\alpha)$ bounded.*

Proof. We have the bound

$$\psi_{\mu L,scaled}(\alpha, y_\delta) \le \psi_{\mu L,scaled}(\alpha, y_\delta - y) + \|R_\alpha(y_\delta - y)\|^{\frac{\mu}{1+\mu}} C_{1,\delta} + C_{2,\delta},$$

where $C_{1,\delta}, C_{2,\delta}$ denote constants, independent of α. Suppose that (3.16) holds for a data error $y_\delta - y \in \mathcal{N}_z \notin D(A^\dagger)$. Then it is well known that $\lim_{\alpha\to 0} \|R_\alpha(y_\delta - y)\| = \infty$, hence, for sufficiently small α, $\|R_\alpha(y_\delta - y)\| \neq 0$ and we get

$$0 < C \le \left(\frac{\int r_\alpha(\lambda)^2 dF_\lambda \|y_\delta - y\|^2}{\int \lambda g_\alpha(\lambda)^2 dF_\lambda \|y_\delta - y\|^2} \right)^{\frac{1}{2+2\mu}} + \frac{C_{1,\delta}}{\|R_\alpha(y_\delta - y)\|^{\frac{1}{1+\mu}}} + \frac{C_{2,\delta}}{\|R_\alpha(y_\delta - y)\|}.$$

Taking the limit and using $y_\delta - y \notin D(A^\dagger)$, we obtain that

$$0 < C \le \liminf_{\alpha\to 0} \left(\frac{\int r_\alpha(\lambda)^2 dF_\lambda \|y_\delta - y\|^2}{\int \lambda g_\alpha(\lambda)^2 dF_\lambda \|y_\delta - y\|^2} \right)^{\frac{1}{2+2\mu}}.$$

However, the numerator in this fraction is bounded (and even tends to 0), and the denominator tends to infinity as $\alpha \to 0$; thus, the right-hand side in this inequality tends to 0, which is a contradiction. ∎

This proposition shows that the modified L-curve (at least as stated in this paper) has a serious flaw. We note that it has been suggested [16] to let μ depend on the regularization parameter as well which could be a way to avoid a negative result like Proposition 5.5.

Another ad-hoc suggestion for a repair of the L-curve method would be to compensate with a negative power of α and use, e.g., $\psi(\alpha, y_\delta) = \frac{1}{\sqrt{\alpha}} \psi_{\mu L}(\alpha, y_\delta)$.

6. Discussion. We have established a rather general framework for the convergence analysis of minimization-based noise level-free parameter choice rules. It is not possible in one paper to cover all cases in detail, but the methodology in this work can, of course, be applied to other situations.

The assumptions in this paper, the noise conditions (4.6), (4.19) and the decay conditions (4.16)–(4.18) are the main tools for the proofs. This type of conditions were first used in [14] and later in [20]. Other authors have used different conditions for a convergence proof, e.g., lower and upper bounds on the Fourier coefficients of the noise: if (σ_i, u_i, v_i) denotes the singular system of a compact operator, then inequalities of the form $c_1 n^{-p} \le (y_\delta - y, v_n) \le c_2 n^{-p}$ can be of value; see, e.g., [17, 18]. Assuming such inequalities and a certain decay of the singular values, (4.6) can be verified [14]. However, such decay rate conditions are not as general as the noise conditions in this paper; moreover, the rate conditions usually need multiple constants and parameters. In this sense, the conditions in this paper are more economical. The same holds for the conditions on $x^\dagger$, where a certain fixed decay rate of its Fourier coefficients usually implies the decay conditions. We note that in [2] scaling conditions were used in place of the noise conditions. It can be shown that the noise condition (4.6) can be rephrased into such scaling condition.

Of course, there is no way to verify the noise condition and decay conditions in reality, unless the noise and the exact solution are known. Concerning the noise condition, this is similar to the usual parameter choice rule, where one assumes the noise level to be known. In noise level-free parameter choice rules, we replace the assumption of a known δ by a qualitative assumption on the noise. The interpretation of the noise condition is that the noise has to be sufficiently "nonsmooth", or, more precisely, it should not be in in $R(A)$. In fact, this agrees with the Bakushinskii veto, where a counterexample for the convergence of noise level-free parameter choice rules is constructed by using noise that is in $R(A)$. The noise condition is usually satisfied for random noise; however, it should be noted that (4.6) also depends on the smoothing properties of A. The more smoothing (in terms of the decay of the singular values) the operator is, the less restrictive is the noise condition. Again, this fits our interpretation that the noise should not be in $R(A)$, because for a highly smoothing operator it is quite unlikely that noise is smooth enough to be close to $R(A)$. The noise condition is central to the convergence of parameter choice rules, thus the least desirable setup for noise level-free parameter choice rule is an operator that is only mildly ill-posed together with data noise that is smooth.

Without additional conditions on the exact solution, the noise level-free parameter choice rules only yield suboptimal convergence rates, unless decay conditions (4.16) or (4.17) hold. These are assumptions on the exact solution that have to be made similar to (and together with) a source condition. While a source condition is a smoothness condition stated independently of the regularization method, the interpretation of the decay condition is that the smoothness of $x^\dagger$ should be close to the qualification of the regularization method (or close to $\mu_0 - \frac{1}{2}$ in the case of $\psi_{HR,\infty}$). Note that if the smoothness index coincides with the qualification index then the decay conditions are trivially satisfied. It is an important fact that the convergence rates for noise level-free parameter choice rules depend on the regularization method in a crucial way, which is different from δ-based rules. From the analysis in this paper, we observed that low saturating methods yield better convergence rates in case of Hölder source conditions. The worst scenario for noise level-free parameter choice rules in terms of convergence rates is a regularization method that has no finite qualification index (such as Landweber iteration, or spectral cutoff) together with an exact solution that has low smoothness. At least for $\psi_{HR,\infty}$ this can be seen from the numerical results in [21, Table 34].

Comparing the methods $\psi_{QO}, \psi_{DQO}, \psi_{HR,\tau}$, we have seen that with regard to convergence they behave quite similar. The Hanke-Raus rules require a weaker noise condition, but the difference between the sets $\mathcal{N}_1$ and $\mathcal{N}_2$ is probably not important in practice. In view of convergence rates, we have observed that $\psi_{QO}, \psi_{DQO}, \psi_{HR,\mu_0}$ yield similar convergence rates (in the finite qualification case), while $\psi_{HR,\infty}$ behaves differently. There is not a general answer for which of the rules is better; rather the choice depends on the smoothness of $x^\dagger$. If the smoothness index μ of $x^\dagger$ is close to the qualification, then the first rules are better, while if μ is close to $\mu_0 - \frac{1}{2}$, the rule $\psi_{HR,\infty}$ is better. In this respect the residual-based method $\psi_{HR,\infty}$ behaves similarly to the discrepancy principle, which also does not yield the best convergence rates in the whole range $\mu \leq \mu_0$, but saturates before [6]. The rules ψ_{QO}, ψ_{DQO} behave quite similarly in many respect, which justifies the view of ψ_{DQO} as a discretized version of ψ_{QO}. Moreover, the rule ψ_{DQO} is probably the easiest one to implement.

Let us finally note that our analysis reveals the reason for the reported success of noise level-free parameter choice rules in practice. As far as we know, most of the numerical comparisons of these rules have used random noise and rather regular exact solutions, which means that the noise conditions and decay conditions were satisfied. Moreover, our analysis can also be used to explain and understand "inverse crimes": it has been proposed that a fair comparison of parameter choice rules should not exclusively use random noise, but also, e.g., the error due to different discretizations. This fits into our analysis since random noise is the good case (the noise condition is satisfied), while discretization errors in the data can be instances where the noise condition is not necessarily satisfied.

In conclusion, it can be said that noise level-free parameter choice rules are a useful enrichment to the arsenal of classical parameter choice rules. However, they should be used only with a good understanding of the structure of the noise in a given problem and of the possible limited convergence properties.

Acknowledgment. The author would like to thank Andreas Neubauer for valuable discussions and advice.

REFERENCES

[1] A. BAKUSHINSKY, *Remarks on choosing a regularization parameter using the quasioptimality and ratio criterion*, U.S.S.R. Comput. Math. Math. Phys., 24 (1984), pp. 181–182.

[2] F. BAUER AND S. KINDERMANN, *The quasi-optimality criterion for classical inverse problems*, Inverse Problems, 24 (2008). 035002 (20 pages).

[3] F. BAUER AND S. KINDERMANN, *Recent results on the quasi-optimality principle*, J. Inverse Ill-posed Probl., 17 (2009), pp. 1129–1142.

[4] F. BAUER AND M. A. LUKAS, *Comparing parameter choice methods for regularization of ill-posed problems*, Math. Comput. Simulation, 81 (2011), pp. 1795–1841.

[5] C. BREZINSKI, G. RODRIGUEZ, AND S. SEATZU, *Error estimates for linear systems with applications to regularization*, Numer. Algorithms, 49 (2008), pp. 85–104.

[6] H. ENGL, M. HANKE, AND A. NEUBAUER, *Regularization of Inverse Problems*, Kluwer, Dordrecht, 1996.

[7] V. GLASKO AND Y. KRIKSIN, *The principle of quasioptimality for linear ill-posed problems in Hilbert space*, Zh. Vychisl. Mat. i Mat. Fiz., 24 (1984), pp. 1603–1613.

[8] U. HÄMARIK, R. PALM, AND T. RAUS, *On minimization strategies for choice of the regularization parameter in ill-posed problems*, Numer. Funct. Anal. Optim., 30 (2009), pp. 924–950.

[9] ———, *Extrapolation of Tikhonov regularization method*, Math. Model. Anal., 15 (2010), pp. 55–68.

[10] M. HANKE AND P. C. HANSEN, *Regularization methods for large-scale problems*, Survey Math. Indust., 3 (1993), pp. 253–315.

[11] M. HANKE AND T. RAUS, *A general heuristic for choosing the regularization parameter in ill-posed problems*, SIAM J. Sci. Comput., 17 (1996), pp. 956–972.

[12] P. C. HANSEN, *Analysis of discrete ill-posed problems by means of the L-curve*, SIAM Rev., 34 (1992), pp. 561–580.

[13] P. C. HANSEN AND D. P. O'LEARY, *The use of the L-curve in the regularization of discrete ill-posed problems*, SIAM J. Sci. Comput., 14 (1993), pp. 1487–1503.

[14] S. KINDERMANN AND A. NEUBAUER, *On the convergence of the quasioptimality criterion for (iterated) Tikhonov regularization*, Inverse Probl. Imaging, 2 (2008), pp. 291–299.

[15] C. L. LAWSON AND R. J. HANSON, *Solving Least Squares Problems*, Prentice-Hall, Englewood Cliffs, NJ, 1974.

[16] S. LU AND P. MATHÉ, *Heuristic parameter selection based on functional minimization: Optimality and model function approach*, Technical Report 1413, WIAS, Berlin, 2009.

[17] M. A. LUKAS, *Asymptotic optimality of generalized cross-validation for choosing the regularization parameter*, Numer. Math., 66 (1993), pp. 41–66.

[18] ———, *Comparisons of parameter choice methods for regularization with discrete noisy data*, Inverse Problems, 14 (1998), pp. 161–184.

[19] P. MATHÉ AND S. PEREVERZEV, *Geometry of linear ill-posed problems in variable Hilbert scales*, Inverse Problems, 19 (2003), pp. 789–803.

[20] A. NEUBAUER, *The convergence of a new heuristic parameter selection criterion for general regularization methods*, Inverse Problems, 24 (2008). 055005 (10 pages).

[21] R. PALM, *Numerical Comparison of Regularization Algorithms for Solving Ill-Posed Problems*, PhD thesis, Institute of Computer Science, University of Tartu, 2010.

[22] T. REGIŃSKA, *A regularization parameter in discrete ill-posed problems*, SIAM J. Sci. Comput., 17 (1996), pp. 740–749.

[23] A. TIKHONOV AND V. GLASKO, *The approximate solution of Fredholm integral equations of the first kind*, Zh. Vychisl. Mat. i Mat. Fiz., 4 (1964), pp. 564–571.

[24] ———, *Use of the regularization method in non-linear problems*, U.S.S.R. Comput. Math. Math. Phys., 5 (1965), pp. 93–107.

[25] G. WAHBA, *Spline Models for Observational Data*, SIAM, Philadelphia, 1990.