Journal of Computational and Applied Mathematics 236 (2012) 2146–2157

Contents lists available at SciVerse ScienceDirect

Journal of Computational and Applied Mathematics

journal homepage: www.elsevier.com/locate/cam

A family of rules for parameter choice in Tikhonov regularization of ill-posed problems with inexact noise level

Uno Hämarik *, Reimo Palm, Toomas Raus

Faculty of Mathematics and Informatics, University of Tartu, J. Liivi 2, 50409 Tartu, Estonia

ARTICLE INFO

Article history:
Received 20 October 2010

MSC:
65J20
47A52

Keywords:
Ill-posed problem
Noise level
Tikhonov regularization
Parameter choice
Discrepancy principle
Monotone error rule

ABSTRACT

We consider Tikhonov regularization of linear ill-posed problems with noisy data. The choice of the regularization parameter by classical rules, such as discrepancy principle, needs exact noise level information: these rules fail in the case of an underestimated noise level and give large error of the regularized solution in the case of very moderate overestimation of the noise level. We propose a general family of parameter choice rules, which includes many known rules and guarantees convergence of approximations. Quasi-optimality is proved for a sub-family of rules. Many rules from this family work well also in the case of many times under- or overestimated noise level. In the case of exact or overestimated noise level we propose to take the regularization parameter as the minimum of parameters from the post-estimated monotone error rule and a certain new rule from the proposed family. The advantages of the new rules are demonstrated in extensive numerical experiments.

© 2011 Elsevier B.V. All rights reserved.

1. Introduction

We consider the operator equation

$$Au = f, \quad f \in \mathcal{R}(A), \tag{1}$$

where $A \in \mathcal{L}(H, F)$ is a linear continuous operator between real Hilbert spaces H and F with norm $\|\cdot\|$ and scalar product $\langle\cdot,\cdot\rangle$. In general our problem is ill-posed: the range $\mathcal{R}(A)$ may be non-closed, the kernel $\mathcal{N}(A)$ may be non-trivial. We suppose that instead of an exact right-hand side f we have only an approximation $\tilde{f} \in F$ and exact or approximative information δ about the noise norm $\|\tilde{f} - f\|$. For most of this paper we assume $\|\tilde{f} - f\| \le \delta$, where δ may also be a serious overestimation, but we also consider the case of an underestimated noise level. In the numerical experiments of Section 5 we increased the rate of over- and underestimation successively by factor 2, using $\varrho := \delta/\|\tilde{f} - f\| = 2^{-6}, 2^{-5}, \dots, 2^5, 2^6$. Note that in applications many sets of data may be available. In the case of two data sets $\tilde{f}_1, \tilde{f}_2$ we may take $\delta := c\|\tilde{f}_1 - \tilde{f}_2\|$ with some constant c.

To find the regularized solution u_α of the equation $Au = f$ we apply the Tikhonov method

$$u_\alpha = (\alpha I + A^*A)^{-1}A^*\tilde{f}, \quad \alpha > 0$$

(see [1,2]). Here an important problem is the choice of a proper regularization parameter α. If α is too small, then the numerical implementation may be unstable; in the case of large α the approximation error is large.

To guarantee convergence of the regularized solutions, the parameter choice must use the noise level (see [3]). Many a posteriori parameter choice rules that use the noise level $\delta \ge \|\tilde{f} - f\|$ can be presented by the following general rule.

* Corresponding author.
E-mail address: uno.hamarik@ut.ee (U. Hämarik).

0377-0427/$ – see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.cam.2011.09.037

General Rule. *Given a function $d(\alpha)$, an estimate δ of the noise level, and a positive parameter b_0, choose the regularization parameter $\alpha(\delta)$ such that*

$$d(\alpha) = b\delta \tag{2}$$

for some parameter $b > b_0$. If the solution does not exist, then choose $\alpha(\delta) = 0$; if this equation has many solutions, choose the smallest solution.

Examples of this general rule are the following rules.

(1) *Discrepancy principle (D rule)* [4,2]. Discrepancy principle has the form of rule (2) with $d(\alpha) = \|Au_\alpha - \tilde{f}\|$ and $b_0 = 1$.

(2) *The modification of discrepancy principle (MD rule)* [5–7]. Here $d(\alpha) = \|B_\alpha(Au_\alpha - \tilde{f})\| = \langle Au_{2,\alpha} - \tilde{f}, Au_\alpha - \tilde{f}\rangle^{1/2}$ and $b_0 = 1$, where the operator B_α is defined by $B_\alpha = \sqrt{\alpha}(\alpha I + AA^*)^{-1/2}$ and $u_{2,\alpha}$ is 2-iterated Tikhonov approximation: $u_{2,\alpha} = (\alpha I + A^*A)^{-1}(\alpha u_\alpha + A^*f)$.

(3) *The monotone error rule (ME rule)* [8,9]. In the case of the ME rule

$$d(\alpha) = \frac{\|B_\alpha(Au_\alpha - \tilde{f})\|^2}{\|B_\alpha^2(Au_\alpha - \tilde{f})\|} = \frac{\langle Au_{2,\alpha} - \tilde{f}, Au_\alpha - \tilde{f}\rangle}{\|Au_{2,\alpha} - \tilde{f}\|}, \qquad b_0 = 1.$$

Due to inequality $\alpha_{\mathrm{ME}} \geq \alpha_{\mathrm{opt}} := \operatorname{argmin}\{\|u_\alpha - u_*\|\}$ [9], typically a better parameter than α_{ME} is the post-estimated parameter $\alpha_{\mathrm{MEe}} = 0.4\alpha_{\mathrm{ME}}$.

Note that in these three rules $d(\alpha)$ is an increasing function of α and Eq. (2) has a unique solution if $\|\tilde{f}\|/b < \delta$. In the case of a very large noise level $\delta > \|\tilde{f}\|/b$ the equations in these rules may have no solutions.

(4) *Rule R*1. In [10,11] an a posteriori parameter choice rule was studied, for which the operators D_α^k were applied to the modified discrepancy, where

$$D_\alpha = \alpha^{-1}AA^*B_\alpha^2.$$

In this rule $d(\alpha) = \|D_\alpha^k B_\alpha(Au_\alpha - \tilde{f})\|$, $2k \in \mathbb{N}$, $k > 0$ and $b_0 = \tilde{\gamma}_k := \frac{(3/2)^{3/2}k^k}{(k+3/2)^{k+3/2}}$. As $B_\alpha^2(Au_\alpha - \tilde{f}) = Au_{2,\alpha} - \tilde{f}$, the norm $\|D_\alpha^k B_\alpha(Au_\alpha - \tilde{f})\|$ can be computed by the formula

$$\|D_\alpha^k B_\alpha(Au_\alpha - \tilde{f})\| = \begin{cases} \alpha^{-k}\langle (AA^*)^k(Au_{k+1,\alpha} - \tilde{f}), (AA^*)^k(Au_{k+2,\alpha} - \tilde{f})\rangle^{1/2}, & \text{if } k \in \mathbb{N}, \\ \alpha^{-k}\|A^*(AA^*)^{k-1/2}(Au_{k+3/2,\alpha} - \tilde{f})\|, & \text{if } k + 1/2 \in \mathbb{N}. \end{cases}$$

(5) *Balancing principle (BP rule)*. This rule has a different form in different papers. Here we consider the form from papers [12–14]:

$$d(\alpha) = \frac{\sqrt{\alpha}\sqrt{q}\|u_\alpha - u_{\alpha/q}\|}{1-q}, \quad 0 < q < 1, \qquad b_0 = \frac{3\sqrt{6}}{16} \approx 0.459.$$

Note that the balancing principle can be considered as an approximation of rule R1 with $k = 1/2$ (see [12]). Typically the balancing principle is implemented by computing a sequence of Tikhonov approximations, but in the case of a smooth solution a better approximation than single Tikhonov approximation is a proper linear combination of Tikhonov approximations with different parameters (see [15]).

(6) *Rule R*2. In [16] a rule was proposed which works better than above mentioned rules in the case of an overestimated noise level:

$$d(\alpha) = d_{\mathrm{R2}}(\alpha) := \frac{\kappa_\alpha^{1/2}\|D_\alpha^{1/2}B_\alpha(Au_\alpha - \tilde{f})\|^2}{\|D_\alpha^{1/2}B_\alpha^2(Au_\alpha - \tilde{f})\|},$$

$\kappa_\alpha := 1 + \alpha\|A\|^{-2}$, $b_0 = \frac{2}{3\sqrt{3}} \approx 0.385$.

The usual way to characterize the quality of various a posteriori parameter choices is to prove their order-optimality on different sets of solutions. The paper [17] introduced the property of quasi-optimality to characterize the quality of an a posteriori regularization parameter choice rule for the particular problem $Au = f$.

Definition 1. A rule R for a posteriori choice of the regularization parameter $\alpha = \alpha(\mathrm{R})$ in the Tikhonov method is *quasi-optimal*, if there exists a constant C (which does not depend on A, u_*, $\tilde{f}$) such that for each $\tilde{f}$, $\|\tilde{f} - f\| \leq \delta$ the error estimate

$$\|u_{\alpha(R)} - u_*\| \leq C \inf_{\alpha \geq 0}\{\|u_\alpha^+ - u_*\| + 0.5\alpha^{-1/2}\delta\}$$

holds. Here u_α^+ is the approximate solution with an exact right-hand side, u_* is the minimum-norm solution of Eq. (1), and the term $0.5\alpha^{-1/2}\delta$ is an upper bound of the norm $\|u_\alpha - u_\alpha^+\|$.

If the Tikhonov method with a priori parameter choice is order-optimal on some set of solutions $M \in H$ and an a posteriori parameter choice rule R is quasi-optimal for this method, then the rule R is order-optimal on the set M also. For example, if

$$M = M_{p,\rho} = \{u \in H : u - u_0 = (A^*A)^{p/2}v,\ \|v\| \le \rho,\ p > 0\},$$

then from quasi-optimality of rule R it follows that

$$\|u_{\alpha(\mathrm{R})-u_*}\| \le C\rho^{\frac{1}{p+1}}\delta^{\frac{p}{p+1}}, \quad 0 < p \le 2.$$

This means that the Tikhonov method with parameter choice by rule R is order-optimal for the full range $p \in (0, 2]$ also.

The rules MD, ME, R1, and BP are quasi-optimal rules for the Tikhonov method, as the following theorem shows.

Theorem 1. *Let $f \in \mathcal{R}(A)$, $\|\tilde{f} - f\| \le \delta$. Let the parameter $\alpha(\delta)$ be chosen according to one of the rules MD, ME, R1, or BP. Then for the Tikhonov method the error estimate*

$$\|u_{\alpha(\delta)} - u_*\| \le \|u^+_{\alpha(\delta)} - u_*\| + \frac{\delta}{2\sqrt{\alpha(\delta)}} \le C(b) \inf_{\alpha \ge 0}\left(\|u^+_\alpha - u_*\| + \frac{\delta}{2\sqrt{\alpha}}\right) \tag{3}$$

holds.

The proof of Theorem 1 is given in [17,12]: the quasi-optimality of rules MD, R1, ME in [17], of rule BP in [12]. Note that the discrepancy principle for the Tikhonov method is not quasi-optimal. For Rule R2 [16] we are able to prove the estimate (3) under a certain additional condition, but extensive numerical experiments indicate the quasi-optimality of this rule, too.

In many papers (cf. [18,19,13,14,20]) for some parameter choice rules the stochastic oracle inequality is proved. Note that quasi-optimality of a rule means that for this rule the deterministic oracle inequality (see [20, Corollary 1]) holds. The first rule for which the quasi-optimality and the oracle inequality were proved is the modified discrepancy principle (with slightly different coefficients) proposed in [7] for the Tikhonov method and in [6] for the Lavrentiev method.

Rules D, MD, and ME are unstable with respect to inaccuracies of the noise level: they fail in the case of an underestimated noise level and give a large error of the Tikhonov approximation already at very moderate overestimation of the noise level. Rules R1, BP and R2 also allow moderate under- or overestimation of the noise level. In rules R1, BP, and R2 the function $d(\alpha)$ is not monotone and the Eq. (2) may have many solutions. In the formulations above of these rules the smallest solution was taken to be the regularization parameter $\alpha(\delta)$; this is good in the case of an exact noise level. In the case of approximate noise level information we recommend using rules R1', BP', and R2', taking the regularization parameter $\alpha(\delta)$ as the largest solution of the corresponding equation (2). In [11] the following theorem was proved for rule R1'.

Theorem 2. *Let $A^*f \in \mathcal{R}(A^*A)$. Let $\alpha = \alpha(\delta)$ be the largest solution of the equation $\|D^k_\alpha B_\alpha(Au_\alpha - \tilde{f})\| = b\delta$, $2k \in \mathbb{N},\ k > 0$. Then under the condition*

$$\lim_{\delta\to 0} \|\tilde{f} - f\|/\delta \le c_*, \quad c_* = \text{const}$$

the convergence

$$\|u_{\alpha(\delta)} - u_*\| \le \|u^+_{\alpha(\delta)} - u_*\| + \frac{\|\tilde{f} - f\|}{2\sqrt{\alpha(\delta)}} \to 0 \quad \text{as } \delta \to 0$$

holds, where u_* is the least square solution of the equation $Au = f$.

In [12], the corresponding theorem was proved for rule BP'. In this paper we consider a general family of rules for a posteriori choice of the regularization parameter, finding the smallest or the largest solution in the corresponding equation, and present convergence results, in the case of the smallest solution also quasi-optimal error estimates. The advantage of the new rules is demonstrated by numerical experiments. In particular, Fig. 3 shows that our rule Me is slightly better than the discrepancy principle in the case of an exact noise level, essentially better in the case of an overestimated noise level, and some proposed new rules allow essential over- and underestimation of the noise level. For easier comparison of the parameter choice rules considered in this paper we note where certain rules are described: known rules D, MD, ME, MEe, R1, BP, R2, R1, BP, R2 in the Introduction (Section 1), known rules DM and QC in Section 5; rules from the general family: rules R(k) in Section 2, rules R(k, β) and R(k, β), Me in Section 5.

Note that results of extensive numerical comparisons of various rules for parameter choice are also presented in [21–24].

2. Motivation of the family of parameter choice rules

In the following we consider a general family of parameter choice rules [25] for the Tikhonov method, which includes several known quasi-optimal parameter choice rules.

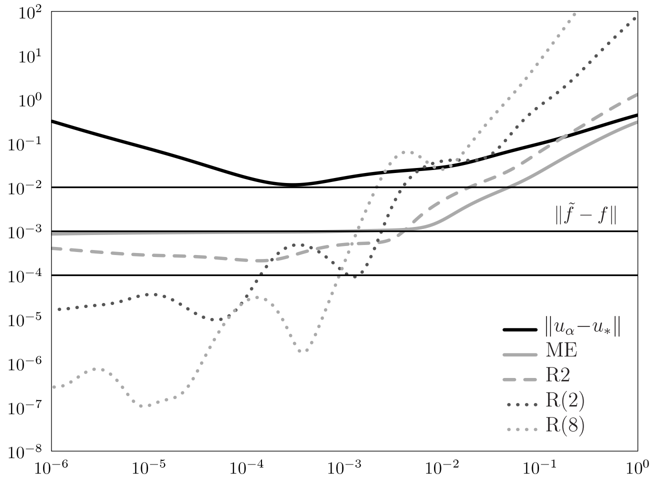

Fig. 1. Behaviours of functions used in different parameter choice rules in the case of a non-smooth solution ($p = 0$) in dependence on α. Three horizontal lines show the exact noise level $\|\tilde{f} - f\|$ and the 10 times over- and underestimated noise levels.

General Rule. *Fix the parameters q, k, s such that $3/2 \le q \le 2, s \ge 0, k \ge s/2$. Let the constant $b > b_0(q, k, s)$. Choose the regularization parameter $\alpha = \alpha(\delta)$ as the solution of the equation*

$$d(\alpha, q, k, s) := \frac{\kappa_\alpha^{2k+s_0-s}\|D_\alpha^k B_\alpha(Au_\alpha - \tilde{f})\|^2}{\|D_\alpha^{sq/2} B_\alpha^{2q-2}(Au_\alpha - \tilde{f})\|^{\frac{2}{q}}} = b\delta^{2-2/q}, \tag{4}$$

where

$$\kappa_\alpha := 1 + \alpha\|A\|^{-2}, \qquad D_\alpha^0 := I, \qquad s_0 := \begin{cases} 0, & \text{if } 2k = s, \\ 1, & \text{if } 2k > s. \end{cases}$$

If the solution does not exist, then choose $u_{\alpha(\delta)} = 0$; if there exists more than one solution, then choose the smallest solution. To avoid computational instabilities when finding the smallest solution of (4), we recommend using some lower bound of α's: either stop before evident numerical instabilities occur or use the lower bound $\underline{\alpha} = (\delta/(2M))^2$, where M is some upper estimate of $\|u_*\|$.

The rule (4) includes as special cases the ME rule ($q = 2, s = k = 0$), the MD rule ($q = 3/2, s = k = 0$), rule R2 ($q = 2, s = 1/2, k = 1/2$), and rule R1 ($q = 3/2, s = 4k/3$). Rules R1 and R2 have slight differences from the corresponding rule (4) in the exponent of κ_α, but due to convergence, $\kappa_\alpha \to 1$ as $\alpha \to 0$, the difference is marginal for small α.

Note that this general family of rules was derived by considering the conditions under which the function

$$t_c(\alpha) = \|u_\alpha^+ - u_*\|^2 + c\|u_\alpha - u_\alpha^+\|^2$$

is monotonically increasing.

In this paper we focus our attention on this family of rules in the case $q = 3/2$. Then we choose the parameter α as the solution of the equation

$$d(\alpha, 3/2, k, s) \equiv \frac{\kappa_\alpha^{2k+s_0-s}\|D_\alpha^k B_\alpha(Au_\alpha - \tilde{f})\|^2}{\|D_\alpha^{3s/4} B_\alpha(Au_\alpha - \tilde{f})\|^{\frac{4}{3}}} = b\delta^{\frac{2}{3}}, \tag{5}$$

which is a special case of (4) for $q = 3/2$. In the numerical experiments of Section 5 we restrict ourselves to the case $s = 2/3$, so that the operator $D_\alpha^{3s/4}$ is easily applicable. For the corresponding rule (5) we use notation R(k).

The stability of a parameter choice with respect to the inaccuracy in noise level information depends on the derivative of the function $d(\alpha)$ in the neighbourhood of $\alpha(\|\tilde{f} - f\|)$: the larger the derivative of $d(\alpha)$ occurring in the neighbourhood of $\alpha(\|\tilde{f} - f\|)$, the more stable the parameter choice rule $d(\alpha) = b\delta$ is with respect to the inaccuracy in noise level.

In the Figs. 1 and 2 we show the difference in behaviour of functions $d(\alpha) = b_0^{-1}d(\alpha, 3/2, k, s)$ in rules R(2) and R(8) from the corresponding functions in the ME rule and rule R2 in the example test problem 'phillips' from [26] (see Section 5) in the case $\|\tilde{f} - f\| = 0.001$. Figs. 1 and 2 correspond to the non-smooth solution u_* and to smoothened solution $(A^*A)^{p/2}u_*$ with $p = 2$, respectively. The horizontal axis shows α's, the dark solid line shows the error $\|u_\alpha - u_*\|$ in the Tikhonov approximation. Note that the functions $d(\alpha)$ in rules R1($k = 1/2$) and R2 have very similar graphs, also the functions in rules D and ME have similar graphs. The parameter α in rule $d(\alpha) = b_0\delta$ is chosen as the intersection of the corresponding graph of $d(\alpha)/b_0$ with level δ: in the case of an exact noise level as the intersection with level $\delta = \|\tilde{f} - f\|$, for example in the

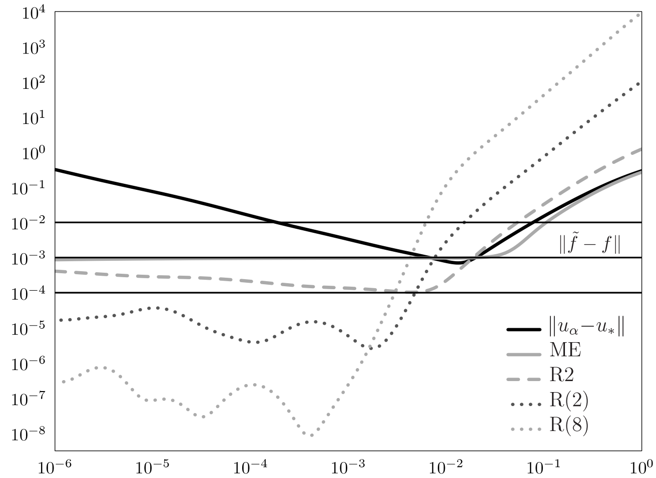

Fig. 2. Behaviours of functions used in different parameter choice rules in the case of a smooth solution ($p = 2$) in dependence on α. Three horizontal lines show the exact noise level $\|\tilde{f} - f\|$ and the 10 times over- and underestimated noise levels.

case of 10 times overestimated or 10 times underestimated noise level we take the intersection with level $\delta = 10\|\tilde{f} - f\|$ or level $\delta = 0.1\|\tilde{f} - f\|$, respectively. If the exact noise level $\delta = \|\tilde{f} - f\|$ is known, many quasi-optimal parameter choices work well. In the case of an inexact noise level δ, different quasi-optimal rules may give very different results. The functions $d(\alpha)$ in rules D and ME are monotone but change only a little in the neighbourhood of $\alpha(\|\tilde{f} - f\|)$, where $\alpha(\|\tilde{f} - f\|)$ is the parameter for which $d(\alpha) = b\|\tilde{f} - f\|$. Therefore small changes of $\delta \approx \|\tilde{f} - f\|$ cause large changes in the corresponding $\alpha(\delta)$'s.

The functions corresponding to rules R(2), R(8) increase much faster in the neighbourhood of $\alpha(\|\tilde{f} - f\|)$ than the function in rule R2 and essentially faster than the functions in rules ME and D. In the case of overestimation of the noise level, using, for example, $\delta = 0.01$, the rules R(2), R(8) work better than rule R2 and the worst result is provided by rule ME. In the case of underestimation of the noise level, using, for example, $\delta = 0.0001$, only rules R(2) and R(8) work comparatively well. Rule R2 does not allow such a large underestimation of the noise level for $p = 0$ and hardly allows it for $p = 2$. Note also that if $\alpha > \alpha(\|\tilde{f} - f\|)$ the function $d(\alpha, q, k, s)$ is monotonically increasing in most cases but if $\alpha \leq \alpha(\|\tilde{f} - f\|)$ the function $d(\alpha, q, k, s)$, $k > s/2$, starts to oscillate. Therefore, in the case of possible overestimation of the noise level it is better to take the regularization parameter as the largest solution of Eq. (4).

3. Limits of the function $d(\alpha, 3/2, k, s)$

The following two lemmas characterize the behaviour of the function $d(\alpha, 3/2, k, s)$. It occurs that this function can be estimated from above by the function used in the quasi-optimal rule R1. If $k \leq 3s/4$ also an upper estimate can be given. In this section we also find the limits of the function $d(\alpha, 3/2, k, s)$ in the processes $\alpha \to 0$ and $\alpha \to \infty$. We use the results of Lemma 1 to prove the quasi-optimality of a sub-family of rules.

Lemma 1. *If $k \geq s/2$, then*

$$d(\alpha, 3/2, k, s) \leq \kappa_\alpha^{2k+s_0-s} \|D_\alpha^{3k-3s/2} B_\alpha (Au_\alpha - \tilde{f})\|^{2/3}.$$

If $k \leq \frac{3s}{4}$, then

$$d(\alpha, 3/2, k, s) \geq \|D_\alpha^k B_\alpha (Au_\alpha - \tilde{f})\|^{2/3}.$$

Proof. To prove the first inequality, we use the inequality of moments $\|Wv\| \leq \|W^p v\|^{1/p} \|v\|^{1-1/p}$ with the operator $W = D_\alpha^{k-3s/4}$ and the element $v = D_\alpha^{3s/4} B_\alpha (Au_\alpha - \tilde{f})$:

$$\|D_\alpha^k B_\alpha (Au_\alpha - \tilde{f})\| \leq \|D_\alpha^{3k-3s/2} B_\alpha (Au_\alpha - \tilde{f})\|^{1/3} \|D_\alpha^{3s/4} B_\alpha (Au_\alpha - \tilde{f})\|^{2/3}.$$

Now we get

$$\begin{aligned} d(\alpha, 3/2, k, s) &= \frac{\kappa_\alpha^{2k+s_0-s} \|D_\alpha^k B_\alpha (Au_\alpha - \tilde{f})\|^2}{\|D_\alpha^{3s/4} B_\alpha (Au_\alpha - \tilde{f})\|^{4/3}} \\ &\leq \kappa_\alpha^{2k+s_0-s} \|D_\alpha^{3k-3s/2} B_\alpha (Au_\alpha - \tilde{f})\|^{2/3}. \end{aligned}$$

To prove the second inequality, we use the inequality $\|D_\alpha^k v\| \le \|v\|$ ($v \in F$, $k \ge 0$), so we get

$$\begin{aligned} d(\alpha, 3/2, k, s) &= \frac{\kappa_\alpha^{2k+s_0-s}\|D_\alpha^k B_\alpha(Au_\alpha - \tilde f)\|^2}{\|D_\alpha^{3s/4}B_\alpha(Au_\alpha - \tilde f)\|^{\frac{4}{3}}} \\ &= \frac{\|D_\alpha^k B_\alpha(Au_\alpha - \tilde f)\|^{4/3}}{\|D_\alpha^{3s/4}B_\alpha(Au_\alpha - \tilde f)\|^{4/3}}\kappa_\alpha^{2k+s_0-s}\|D_\alpha^k B_\alpha(Au_\alpha - \tilde f)\|^{2/3} \\ &\ge \|D_\alpha^k B_\alpha(Au_\alpha - \tilde f)\|^{2/3}. \quad \square \end{aligned}$$

Lemma 2. *For the function $d(\alpha, 3/2, q, s)$ the following limits hold:*

$$\lim_{\alpha\to 0} d(\alpha, 3/2, k, s) = 0, \quad \text{if } 2k > s, \tag{6}$$

$$\lim_{\alpha\to 0} d(\alpha, 3/2, k, s) \le \delta^{2-2/q}, \quad \text{if } 2k = s \text{ and } \|\tilde f - f\| \le \delta, \tag{7}$$

$$\lim_{\alpha\to \infty} d(\alpha, 3/2, k, s) = \infty, \quad \text{if } 2k > s, \tag{8}$$

$$\lim_{\alpha\to \infty} d(\alpha, 3/2, k, s) = \frac{\|(AA^*)^{s/2}\tilde f\|^2}{\|(AA^*)^{3s/4}\tilde f\|^{4/3}}, \quad \text{if } 2k = s. \tag{9}$$

Proof. The following limits are valid [17]:

$$\lim_{\alpha\to 0}\kappa_\alpha^{2k+s_0-s} = 1; \qquad \lim_{\alpha\to 0}\|D_\alpha^{3k-3s/2}B_\alpha(Au_\alpha - \tilde f)\| = 0, \quad \text{if } k > \frac{s}{2};$$
$$\lim_{\alpha\to 0}\|D_\alpha^{3k-3s/2}B_\alpha(Au_\alpha - \tilde f)\| = \lim_{\alpha\to 0}\|B_\alpha(Au_\alpha - \tilde f)\| \le \delta, \quad \text{if } k = \frac{s}{2} \text{ and } \|\tilde f - f\| \le \delta.$$

Therefore the limits (6), (7) follow from the results of Lemma 1.

The function $d(\alpha, 3/2, q, s)$ can be presented in the form

$$d(\alpha, 3/2, k, s) = \frac{\kappa_\alpha^{2k+s_0-s}\alpha^{-2k+s}\|(AA^*)^k B_\alpha^{2k+1}(Au_\alpha - \tilde f)\|^2}{\|(AA^*)^{3s/4}B_\alpha^{3s/2+1}(Au_\alpha - \tilde f)\|^{4/3}}.$$

Consequently from limits

$$\begin{aligned} \lim_{\alpha\to\infty}\alpha^{-2k+s}\kappa_\alpha^{2k+s_0-s} &= \lim_{\alpha\to\infty}(\alpha^{-1} + \|A\|^{-2})^{2k-s}(1+\alpha\|A\|^{-2})^{s_0} \\ &= \begin{cases} 1, & \text{if } 2k = s, \\ \infty, & \text{if } 2k > s \end{cases} \end{aligned}$$

and

$$\lim_{\alpha\to\infty}\|B_\alpha^t(Au_\alpha - \tilde f)\| = \|\tilde f\|, \quad \text{if } t \ge 0$$

we derive the limits (8)–(9). $\square$

From Lemma 2 follows that the Eq. (5) always has a solution if $2k > s$. In the case $2k = s$ this equation has a solution if $\delta \le \|(AA^*)^{s/2}\tilde f\|^3/b^{3/2}\|(AA^*)^{3s/4}\tilde f\|^2$.

4. Convergence and quasi-optimality of family of rules

In this section we show that the choice of the regularization parameter in the Tikhonov method by the rules of the proposed family guarantees both convergence of regularized solutions and a quasi-optimal error estimate for $s/2 \le k \le 3s/4$. It is an open problem whether the quasi-optimality holds also for $k > 3s/4$.

Theorem 3. *Let $\alpha = \alpha(\delta)$ be the smallest solution of the equation $d(\alpha, 3/2, k, s) = b\delta^{2/3}$, $k \ge s/2$, $s \ge 0$,*

$$b > b_{k,s} \equiv \frac{(2k-s)^{2k-s}}{(2k-s+1)^{2k-s+1}}. \tag{10}$$

If $\|\tilde f - f\| \le \delta$, then

$$\|u_{\alpha(\delta)} - u_*\| \to 0 \quad \text{as } \delta \to 0.$$

Proof. It is easy to show that

$$\|u_{\alpha(\delta)} - u_*\| \le \|u^+_{\alpha(\delta)} - u_*\| + \frac{\delta}{2\sqrt{\alpha(\delta)}}, \tag{11}$$

$$Au_\alpha - \tilde{f} = -B^2_\alpha Au_* - B^2_\alpha(\tilde{f} - f),$$

$$\begin{aligned} \|D^k_\alpha B^3_\alpha Au_*\| &\le \|D^k_\alpha B_\alpha(Au_\alpha - \tilde{f})\| + \|D^k_\alpha B^3_\alpha(\tilde{f} - f)\| \\ &\le \|D^k_\alpha B_\alpha(Au_\alpha - \tilde{f})\| + \delta. \end{aligned} \tag{12}$$

If $\delta \to 0$, then the equality $d(\alpha(\delta), 3/2, k, s) = b\delta^{2/3}$ leads to

$$\|D^k_{\alpha(\delta)} B_{\alpha(\delta)}(Au_{\alpha(\delta)} - \tilde{f})\| \to 0$$

and therefore using the inequality (12) we get

$$\|D^k_{\alpha(\delta)} B^3_{\alpha(\delta)} Au_*\| \to 0, \quad \text{as } \delta \to 0.$$

Denote $\bar{B}_\alpha = \sqrt{\alpha}(\alpha I + A^*A)^{-1/2}$. If in the process $\|D^k_{\alpha(\delta)} B^3_{\alpha(\delta)} Au_*\| \to 0$ the parameter $\alpha(\delta) \to 0$, then

$$\|u^+_{\alpha(\delta)} - u_*\| = \|\bar{B}^2_{\alpha(\delta)} u_*\| \to 0.$$

Consider now the case, if in the process $\|D^k_{\alpha(\delta)} B^3_{\alpha(\delta)} Au_*\| \to 0$ the parameter $\alpha(\delta) \ge \bar{\alpha} = \text{const}$. In [2, p. 43,66] it is proved that if $A\bar{B}^2_{\alpha_n} u_* \to 0$ $(n \to \infty)$, $\alpha_n \ge \bar{\alpha} = \text{const}$, then $\bar{B}^2_{\alpha_n} u_* \to 0$ $(n \to \infty)$. Similarly we can show that if $D^k_{\alpha_n} B^3_{\alpha_n} Au_* = \alpha_n^{-k} A(A^*A)^k \bar{B}^{2k+3}_{\alpha_n} u_* \to 0$ $(n \to \infty)$, then $\bar{B}^2_{\alpha_n} u_* \to 0$ $(n \to \infty)$. Hence, $\|u^+_{\alpha(\delta)} - u_*\| \to 0$ $(\delta \to 0)$ and the convergence of the first term of (11) is proved.

To prove the convergence of the second term of (11) we use Lemma 1. Let the parameter $\alpha_1(\delta)$ be chosen by rule R1 with $k' = 3k - 3s/2$, hence the parameter $\alpha_1(\delta)$ is the minimal solution of the equation $\|D^{3k-3s/2}_\alpha B_\alpha(Au_\alpha - \tilde{f})\| = c\tilde{\gamma}_{3k-3s/2}\delta$, $1 < c < (b/b_{k,s})^{3/2}$, $\tilde{\gamma}_{3k-3s/2} = \left[\frac{(2k-s)^{2k-s}}{(2k-s+1)^{2k-s+1}}\right]^{\frac{3}{2}}$. Because of the quasi-optimality of Rule R1, we have

$$\frac{\delta}{\sqrt{\alpha_1(\delta)}} \to 0 \quad \text{as } \delta \to 0. \tag{13}$$

If $\alpha(\delta) \ge \bar{\alpha} > 0$ in the process $\delta \to 0$, then the convergence of the second term of (11) is obvious. If $\alpha(\delta) \to 0$ as $\delta \to 0$, then using Lemma 1, we have for sufficiently small δ for every $\alpha \le \alpha_1(\delta)$

$$\begin{aligned} d(\alpha, 3/2, k, s) &\le \kappa_\alpha^{2k+s_0-s} \|D^{3k-3s/2}_\alpha B_\alpha(Au_\alpha - \tilde{f})\|^{2/3} \\ &= \kappa_\alpha^{2k+s_0-s} c^{2/3} b_{k,s} \delta^{2/3} \le b\delta^{2/3}. \end{aligned}$$

As $\alpha(\delta)$ is the minimal solution of the equation $d(\alpha, 3/2, k, s) = b\delta^{2/3}$, it follows from the last inequality that $\alpha(\delta) \ge \alpha_1(\delta)$, which with (13) proves the convergence of the second term of (11). □

Theorem 4. *Let $\alpha = \alpha(\delta)$ be the largest solution of the equation $d(\alpha, 3/2, k, s) = b\delta^{2/3}$, $b = \text{const}$, $k > s/2$, $s \ge 0$. If in the process $\delta \to 0$ $\frac{\|\tilde{f}-f\|}{\delta} \le c_* = \text{const}$, then*

$$\|u_{\alpha(\delta)} - u_*\| \to 0 \quad \textit{as } \delta \to 0.$$

Proof. The following inequality holds

$$\|u_{\alpha(\delta)} - u_*\| \le \|u^+_{\alpha(\delta)} - u_*\| + \frac{\|\tilde{f} - f\|}{2\sqrt{\alpha(\delta)}}. \tag{14}$$

The proof of the convergence of the first term of (14) is the same as in Theorem 3. In the following we prove the convergence of the second term of (14). Let the parameter $\alpha_1(\delta)$ be the largest solution of the equation $\|D^{3k-3s/2}_\alpha B_\alpha(Au_\alpha - \tilde{f})\| = b'\delta$, $b' < b^{3/2}$. Taking into account that rule R1' is a stable parameter choice rule (see Theorem 2), in the process $\delta \to 0$ we have

$$\frac{\|\tilde{f} - f\|}{\sqrt{\alpha_1(\delta)}} \to 0 \quad \text{if } \frac{\|\tilde{f} - f\|}{\delta} \le c_*. \tag{15}$$

Using Lemma 1 we get for δ small enough

$$\begin{aligned} d(\alpha_1(\delta), 3/2, k, s) &\le \kappa_{\alpha_1(\delta)}^{2k+s_0-s} \|D^{3k-3s/2}_{\alpha_1(\delta)} B_{\alpha_1(\delta)}(Au_{\alpha_1(\delta)} - \tilde{f})\|^{2/3} \\ &= \kappa_{\alpha_1(\delta)}^{2k+s_0-s} b'^{2/3} \delta^{2/3} \le b\delta^{2/3}. \end{aligned}$$

As $\lim_{\alpha\to\infty} d(\alpha, 3/2, k, s) = \infty$ provided that $k > s/2$, for the largest solution $\alpha = \alpha(\delta)$ of the equation $d(\alpha, 3/2, k, s) = b\delta^{2/3}$ the inequality $\alpha(\delta) \geq \alpha_1(\delta)$ holds, which with (15) proves the convergence of the second term of (14). □

Theorem 4 shows that the rules with $k > s/2,\ s \geq 0$ are stable parameter choice rules, if we choose the largest solution of the equation $d(\alpha, 3/2, k, s) = b\delta^{2/3}$. This means that the error of the approximate solution converges to zero also in the case where the noise level of the right-hand side is greater than δ.

Theorem 5. *Let $\|\tilde{f} - f\| \leq \delta$. Let $\alpha = \alpha(\delta)$ be the minimal solution of the equation $d(\alpha, 3/2, k, s) = b_0\delta^{2/3},\ b_0 > \tilde{\gamma}_k, s/2 \leq k \leq 3s/4,\ s \geq 0$. If $\alpha(\delta) < \|A\|^2[(b_0/b_{k,s})^{1/(2k-s_0+s)} - 1]$ then*

$$\|u_{\alpha(\delta)} - u_*\| \leq C(b_0) \inf_{\alpha\geq 0} \left\{ \|u_\alpha^+ - u_*\| + \frac{\delta}{2\sqrt{\alpha}} \right\}.$$

Proof. Using Lemma 1, we get

$$b_0\delta^{2/3} = d(\alpha(\delta), 3/2, k, s) \leq \|D_{\alpha(\delta)}^{3k-3s/2} B_{\alpha(\delta)}(Au_{\alpha(\delta)} - \tilde{f})\|^{2/3} \tag{16}$$

and

$$b_0\delta^{2/3} \geq d(\alpha, 3/2, k, s) \geq \kappa_\alpha^{2k-s_0+s} \|D_\alpha^k B_\alpha(Au_\alpha - \tilde{f})\|^{2/3}, \quad \alpha \leq \alpha(\delta). \tag{17}$$

Let the parameter $\alpha_1(\delta)$ be chosen by rule R1 with $k' = 3k - 3s/2$ and $b = b_0/\kappa_\alpha^{2k-s_0+s} > \tilde{\gamma}_{3k-3s/2}$ and let the parameter $\alpha_2(\delta)$ be chosen by rule R1 with $k' = k$ and $b = {b_0}^{3/2}$. From (16), (17) it follows that $\alpha_1(\delta) \leq \alpha(\delta) \leq \alpha_2(\delta)$. Since Rule R1 is a quasi-optimal parameter choice rule, the inequality

$$\|u_{\alpha_j(\delta)}^+ - u_*\| + \frac{\delta}{2\sqrt{\alpha_j(\delta)}} \leq C_j(b_0) \inf_{\alpha\geq 0} \left\{ \|u_\alpha^+ - u_*\| + \frac{\delta}{2\sqrt{\alpha}} \right\}, \quad j = 1, 2$$

holds true (see Theorem 1). Since the function $t(\alpha) = \|u_\alpha^+ - u_*\|$ is monotonically increasing, we get

$$\begin{aligned} \|u_{\alpha(\delta)} - u_*\| &\leq \|u_{\alpha(\delta)}^+ - u_*\| + \frac{\delta}{2\sqrt{\alpha(\delta)}} \leq \|u_{\alpha_2(\delta)}^+ - u_*\| + \frac{\delta}{2\sqrt{\alpha_1(\delta)}} \\ &\leq C(b_0) \inf_{\alpha\geq 0} \left\{ \|u_\alpha^+ - u_*\| + \frac{\delta}{2\sqrt{\alpha}} \right\}, \end{aligned}$$

which proves Theorem 5. □

Consider the choice of the coefficient b in Eq. (5). If the smallest solution is used we recommend the coefficient b from condition (10). If the largest solution is used, it is reasonable to take b smaller.

Under the coarse assumption that different quasi-optimal parameter choices give close parameters, there exists a parameter α_0 for which

$$d(\alpha_0, 2/3, k, s) = \frac{\kappa_{\alpha_0}^{2k+s_0-s} \|D_{\alpha_0}^k B_{\alpha_0}(Au_{\alpha_0} - \tilde{f})\|^2}{\|D_{\alpha_0}^{3s/4} B_{\alpha_0}(Au_{\alpha_0} - \tilde{f})\|^{\frac{4}{3}}} \approx \frac{\tilde{\gamma}_k^2}{(\tilde{\gamma}_{3s/4})^{4/3}} \delta^{2/3}$$

and therefore the coefficient b can be chosen to be equal to

$$\tilde{b}_{k,s} = \frac{\tilde{\gamma}_k^2}{(\tilde{\gamma}_{3s/4})^{4/3}} = \frac{3}{2} \frac{k^{2k}}{(3s/4)^s} \frac{(3s/4 + 3/2)^{s+2}}{(k + 3/2)^{2k+3}}.$$

Numerical results show that if for the regularization parameter the largest solution of Eq. (5) is used, then it is reasonable to use b near $\tilde{b}_{k,s}$ for non-smooth solutions and b near $b_{k,s}$ for smooth solutions.

5. Numerical results

Our tests are performed on the well-known set of test problems in [26]: *baart*, *deriv2*, *foxgood*, *gravity*, *heat*, *ilaplace*, *phillips*, *shaw*, *spikes*, *wing* (problems 1–10). In addition, we used test problems from [27]: *gauss*, *hilbert*, *lotkin*, *moler*, *pascal*, *prolate* (problems 11–16). As in [27], we combined these 6 $N \times N$ matrices with 6 solution vectors $u_i = 1$, $u_i = i/N$, $u_i = ((i - \lfloor N/2 \rfloor)/\lceil N/2 \rceil)^2$, $u_i = \sin(2\pi(i-1)/N)$, $u_i = i/N + 1/4\sin(2\pi(i-1)/N)$, $u_i = 0$ if $i \leq \lfloor N/2 \rfloor$ and 1 if $i > \lfloor N/2 \rfloor$. In all tests we had $N = 100$. In Table 1 the upper part corresponds to problems 1–10 and the lower part to problems 11–16.

Since the performance of methods and rules generally depends on the smoothness p of the exact solution, we complemented the standard solutions u_* of (now discrete) test problems with smoothened solutions $(A^*A)^{p/2}u_*$ $(p = 2)$ (computing the right-hand side as $A\big((A^*A)^{p/2}u_*\big)$). After discretization all problems were scaled (normalized) in such a way that the Euclidean norms of the operator and the right-hand side were 1. On a base of exact data f we formed the noisy data $\tilde{f}$, where $\|\tilde{f} - f\|$ had values 0.3, 10^{-1}, 10^{-2}, 10^{-3}, 10^{-4}, 10^{-5}, 10^{-6}. The noise $\tilde{f} - f$ added to f had a normal distribution. Here, $\|\cdot\|$ means Euclidean norm.

Table 1
Averages of error ratios in problems in the case $p = 0$, $\varrho = 1$. (a) % of cases where R(2, 1) ≡ R'(2, 1); (b) % of cases where R(8, 1) ≡ R'(8, 1).

	D	MEe	R2'	R(2, 1)	R'(2, 0.25)	R'(2, 0.5)	(a)	R(8, 1)	R'(8, 0.1)	R'(8, 0.5)	(b)
(1) Baart	1.30	1.28	1.61	1.67	1.47	1.95	50.0	1.70	1.67	2.10	20.0
(2) Deriv2	1.23	1.10	1.14	1.16	1.22	1.05	100	1.30	1.03	1.10	100
(3) Foxgood	1.48	1.82	4.62	4.15	2.25	3.45	100	3.43	1.83	4.54	71.4
(4) Gravity	1.17	1.15	1.48	1.40	1.11	1.11	100	1.41	1.19	1.15	100
(5) Heat	1.07	1.04	1.10	1.18	1.23	1.06	100	1.27	1.23	1.11	85.7
(6) i_laplace	1.27	1.23	1.42	1.43	1.16	1.23	100	1.62	1.14	1.40	100
(7) Phillips	1.04	1.03	1.23	1.16	1.21	1.05	100	1.07	1.45	1.07	95.7
(8) Shaw	1.23	1.20	1.42	1.37	1.32	1.37	71.4	1.44	1.36	1.63	47.1
(9) Spikes	1.01	1.01	1.02	1.02	1.02	1.03	28.6	1.02	1.03	1.04	30.0
(10) Wing	1.15	1.14	1.34	1.28	1.39	1.44	57.1	1.27	1.40	1.46	28.6
Avg 1–10 ($p = 0$)	1.20	1.20	1.64	1.58	1.34	1.47	80.7	1.55	1.33	1.66	67.9
(11) Gauss	1.30	1.26	1.53	1.51	1.16	1.23	92.4	1.50	1.19	1.34	85.7
(12) Hilbert	1.35	1.28	1.68	1.70	1.29	1.53	77.1	1.79	1.31	1.90	53.6
(13) Lotkin	1.65	1.40	1.88	1.88	1.50	1.79	67.1	1.97	1.96	2.85	61.7
(14) Moler	1.06	1.05	1.23	1.24	1.41	1.25	92.9	1.36	1.54	1.99	89.3
(15) Pascal	1.02	1.02	1.04	1.03	1.06	1.06	42.9	1.03	1.06	1.06	42.9
(16) Prolate	1.69	1.78	1.72	1.69	1.52	1.83	94.3	1.71	1.60	2.07	74.0
Avg 11–16 ($p = 0$)	1.34	1.30	1.51	1.51	1.32	1.45	77.8	1.56	1.44	1.87	67.9
Avg 1–10 ($p = 2$)	2.70	1.12	1.23	1.08	1.32	1.14	100	1.18	1.89	1.40	100
Avg 11–16 ($p = 2$)	2.69	1.29	2.18	1.71	1.39	1.37	100	1.31	1.98	1.55	99.2

Table 2
Averages of error ratios over all problems in the case of an overestimated noise level ϱ. (a) % of cases where R(2, 1) ≡ R'(2, 1); (b) % of cases where R(8,1) ≡ R'(8, 1).

	Case $p = 0$, values of ϱ							Case $p = 2$, values of ϱ						
	1	1.1	1.3	2	4	16	64	1	1.1	1.3	2	4	16	64
R2'	1.59	1.63	1.69	1.86	2.30	3.56	5.18	1.58	1.66	1.83	2.38	3.84	8.93	16.0
R(2, 1)	1.55	1.58	1.62	1.75	2.19	3.05	4.56	1.31	1.33	1.38	1.62	2.25	5.17	11.7
R'(2, 0.25)	1.33	1.32	1.34	1.42	1.51	2.01	2.64	1.35	1.33	1.30	1.25	1.21	1.55	3.12
R'(2, 0.5)	1.47	1.48	1.50	1.57	1.78	2.46	3.68	1.23	1.23	1.22	1.23	1.36	2.57	5.90
(a)	78.4	78.7	77.3	74.9	79.3	83.2	86.1	100	100	100	99.7	99.7	100	99.7
R(8, 1)	1.56	1.58	1.63	1.76	2.13	2.89	4.13	1.23	1.22	1.21	1.19	1.20	1.69	3.48
R'(8, 0.05)	1.44	1.43	1.42	1.38	1.38	1.57	1.93	2.20	2.17	2.13	2.01	1.84	1.58	1.34
R'(8, 0.1)	1.38	1.38	1.38	1.40	1.48	1.82	2.34	1.93	1.90	1.86	1.77	1.64	1.40	1.33
R'(8, 0.5)	1.74	1.76	1.79	1.93	2.19	2.94	4.22	1.45	1.44	1.41	1.35	1.35	1.35	2.40
(b)	67.9	67.8	67.4	67.7	71.9	74.1	81.4	99.4	99.4	99.4	99.4	99.4	99.4	99.2

We generated 10 noise vectors and used these vectors in all problems (we also have partially computed with 100 noise vectors but the results were almost the same). The problems were regularized by the Tikhonov method, choosing the regularization parameters by the rules that we wanted to compare. In our experiments we also took into account the possibility of over- or underestimation of the noise level up to several tens of times.

Since in model equations the exact solution is known, it is possible to find the regularization parameter $\alpha = \alpha_*$ which gives the smallest error: $\|u_{\alpha_*} - u_*\| = \min_{\alpha>0}\{\|u_\alpha - u_*\|\}$. For every rule R the error ratio $\|u_{\alpha_R} - u_*\|/\|u_{\alpha_*} - u_*\|$ describes the performance of rule R on this particular problem. To compare the rules or to present their properties, the following tables show averages of these error ratios over various parameters of the data set (problems 1–16, smoothness indices p, noise levels δ, runs). We denote by R(k, β) and R'(k, β), rules which choose the parameter α as the smallest or largest solutions, respectively, of Eq. (5) with $s = 2/3$ and $b = \beta b_{k,2/3}$. To avoid computational instabilities in rules R(2, β) and R(8, β), we used a lower bound for the α's: if at some $\alpha = \underline{\alpha}$ division by zero occurred, then the lower bound of the α's was taken to be $400\underline{\alpha}$. This lower bound was needed only for rules R(2, 1) and R(8, 1).

Table 1 shows averages of error ratios if the exact noise level $\delta = \|\tilde{f} - f\|$ was used. Results are given in the case $p = 0$ for every problem 1–16 and in the case $p = 2$ as averages over problems 1–10 and 11–16. Best results were obtained in the case $p = 0$ in rules MEe, D and in new rules R'(2, 0.25), R'(8, 0.1); in the case $p = 2$ the best results were in rule MEe and in rules R'(2, 0.5), R(8, 1). The discrepancy principle gives large errors in the case $p = 2$ due to saturation at $p = 1$. We conclude that for smoother solutions a larger coefficient β is preferable in the new rules. Columns (a) and (b) show in what percentage of cases Eq. (5) with $b = b_{k,2/3}$ had a unique solution. It turned out in our numerical experiments that this equation had a unique solution in most cases for $k = 8$, and also for $k = 2$ even in most of the problems of [26] (and always for $p = 2$).

Tables 2 and 3 contain results for the cases of overestimated noise level (case $\varrho := \delta/\|\tilde{f} - f\| \geq 1$) and underestimated noise level (case $\varrho \leq 1$), respectively. Table 2 shows that the new rules with $k = 2$ and $k = 8$ allow large overestimation of the noise level. Table 3 shows that rules R'(2, 0.5) and R'(8, 0.5) allow mild and large underestimation of the noise level,

Table 3
Averages of error ratios over all problems in the case of an underestimated noise level ϱ.

	Case $p = 0$, values of ϱ							Case $p = 2$, values of ϱ						
	1/64	1/16	1/4	1/2	0.75	0.9	1	1/64	1/16	1/4	1/2	0.75	0.9	1
R2'	1e13	2e12	848	3.42	1.55	1.56	1.59	2e13	4e12	2e3	2.95	1.40	1.51	1.58
R'(2, 0.25)	4e7	1e3	7.62	1.46	1.36	1.34	1.33	9e7	4e3	7.81	1.58	1.42	1.38	1.35
R'(2, 0.5)	1e4	24.5	1.37	1.37	1.43	1.45	1.47	7e3	21.7	1.44	1.30	1.25	1.23	1.23
R'(8, 0.05)	823	25.3	1.94	1.58	1.50	1.46	1.44	3e3	24.5	2.78	2.44	2.29	2.24	2.20
R'(8, 0.1)	77.1	2.20	1.52	1.41	1.38	1.38	1.38	57.2	3.02	2.31	2.10	2.00	1.95	1.93
R'(8, 0.5)	1.58	1.38	1.45	1.56	1.69	1.72	1.74	2.44	2.02	1.70	1.58	1.50	1.47	1.45

Table 4
Averages of error ratios over all problems for combined rules. Rows: (a) D; (b) $\alpha = 0.9\min(\alpha_{\mathrm{ME}}, \alpha_{\mathrm{R2'}}, \alpha_{\mathrm{R'(2,0.3)}})$; (c) $\alpha = 0.5\min(\alpha_{\mathrm{ME}}, \alpha_{\mathrm{R2'}}, \alpha_{\mathrm{R'(8,1)}})$; (d) $\alpha = \min(0.4\alpha_{\mathrm{ME}}, 1.4\alpha_{\mathrm{R'(2,0.25)}})$; (e) $\alpha = \min(0.5\alpha_{\mathrm{ME}}, 0.5\alpha_{\mathrm{R2'}}, 1.4\alpha_{\mathrm{R'(2,0.25)}})$; (f) $\alpha = \min(0.4\alpha_{\mathrm{ME}}, 5\alpha_{\mathrm{R'(8,0.05)}})$; (g) $\alpha = \min(0.45\alpha_{\mathrm{ME}}, 1.5\alpha_{\mathrm{R'(2,0.25)}}, 3.4\alpha_{\mathrm{R'(8,0.04)}})$.

	Case $p = 0$, values of ϱ							Case $p = 2$, values of ϱ						
	1	1.1	1.3	2	4	16	64	1	1.1	1.3	2	4	16	64
(a)	1.25	1.58	1.83	2.33	2.96	4.28	6.60	2.70	2.70	3.34	5.43	7.57	12.9	28.2
(b)	1.22	1.32	1.34	1.39	1.49	2.05	2.75	1.34	1.34	1.31	1.27	1.23	1.62	3.32
(c)	1.24	1.35	1.38	1.46	1.71	2.30	3.23	1.76	1.74	1.70	1.60	1.46	1.34	1.84
(d)	1.23	1.29	1.31	1.43	1.55	2.13	2.83	1.25	1.21	1.19	1.19	1.24	1.95	4.23
(e)	1.22	1.30	1.31	1.42	1.54	2.12	2.82	1.23	1.21	1.19	1.18	1.23	1.92	4.10
(f)	1.20	1.31	1.33	1.45	1.56	1.89	2.48	1.21	1.16	1.16	1.15	1.16	1.37	1.81
(g)	1.22	1.27	1.28	1.32	1.44	1.66	2.21	1.33	1.30	1.28	1.23	1.17	1.15	1.30

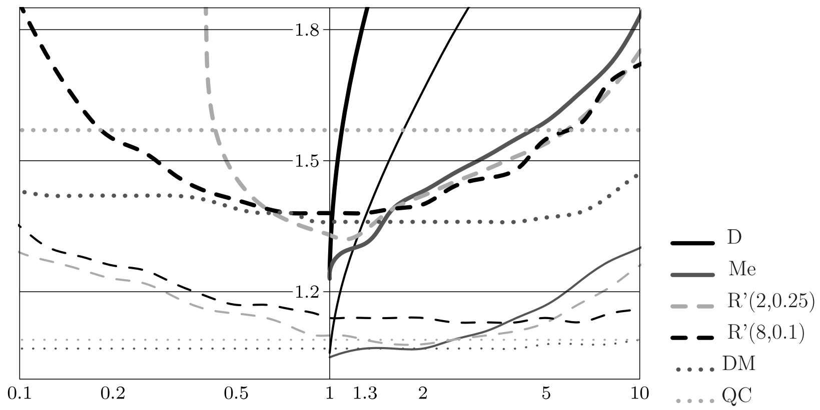

Fig. 3. Averages (thick lines) and medians (thin lines) of error ratios of different parameter choice rules in dependence on $\varrho = \delta/\|\tilde{f} - f\|$.

respectively. Table 3 also confirms that for a smoother solution a larger coefficient β is good. Although the tables do not contain results for classical rules, note that classical parameter choice rules as a discrepancy principle fail in the case of underestimation of the noise level and give a large error in the case of overestimation of the noise level.

Table 4 shows that in the case of exact or overestimated noise level it is good to choose the regularization parameter as the minimum of the parameter α_{MEe} and one or more parameters of the form $c\alpha_{\mathrm{R}}$, where α_{R} is parameter from rule R'(2, β) or R'(8, β) with some $\beta \leq 1$. The coefficient(s) c here is/are some trade-off: for larger overestimation smaller c is good, but for smoother solutions larger c is preferred.

Fig. 3 compares error ratios of different parameter choices with respect to the accuracy $\varrho = \delta/\|\tilde{f} - f\|$ of the noise level in the interval $\varrho \in [0.1, 10]$ in the case $p = 0$. Here thick and thin lines show averages and medians, respectively, of error ratios for corresponding parameter choices. Notation Me is used for rule (d) from Table 4. Note that rules from rows (e)–(g) gave better results but these rules need more computations. Fig. 3 shows that the Me rule typically gives slightly higher accuracy than the discrepancy principle in the case of an exact noise level but much higher accuracy in the case of an overestimated noise level. In the case of an underestimated noise level $\varrho < 1$ both these rules fail, but rules R'(2, 0.25) and R'(8, 0.1) allow underestimation, rule R'(8, 0.1) even allows large underestimation.

Fig. 3 also contains results for rule QC (see [22,24,23]) not using noise level information and for rule DM (see [28,23]) using an approximate noise level. Rule QC is a modification of the quasi-optimality criterion, which chooses $\alpha = \alpha_{\mathrm{Q}}$ as the minimizer of the function $\psi(\alpha) := \|u_\alpha - u_{2,\alpha}\|$. In some problems (for example, in problem 'heat' of [26]) the quasioptimality criterion fails, choosing too small α [22–24]. Rule QC finds α_{QC} as the minimizer of the function $\psi(\alpha)$ for the decreasing α-sequence (we used $\alpha_n = 0.9^{n-1}$, $n = 1, 2, \ldots$) stopping the calculations if $\psi(\underline{\alpha}) \geq C\psi(\alpha_{QC})$ for some $\underline{\alpha}$. We used the constant $C = 5$. Rule DM finds $\underline{\alpha}$ as the maximal solution of the equation $\psi(\alpha)\alpha^{1/2} = c_1\delta$ and then $\alpha = \alpha_{\mathrm{DM}}$ as the minimizer of $d_{R2}(\alpha)\alpha^{c_2}$ in $[\underline{\alpha}, 1]$. We use the constants $c_1 = 0.005$, $c_2 = 0.45$, if information $\rho \in (0.1, 10)$ is available, and $c_1 = 0.001$, $c_2 = 0.03$, if only information $\rho \in (0.01, 100)$ is available. We have added three horizontal level lines at 1.2, 1.5, 1.8, to enable us to estimate at which ϱ this ratio of error level is achieved. Note that for $\varrho < 0.5$ the averages of the error ratios in rule R'(2,0.25) are much larger than in rule R'(8, 0.1) but the medians for $\varrho \in [0.1, 5]$ are less. Fig. 3 shows that rules R'(2, 0.25), R'(8, 0.1) are good for such inaccurate information about the noise level as $\varrho \in [0.1, 10]$.

On the basis of Fig. 3 we recommend using rule Me in the case $\varrho := \delta/\|\tilde{f} - f\| \in [1, 1.5]$ and rule R(2) in the case $\varrho \in [0.6, 1.5]$. In the case of less information about the noise level we recommend rule DM in the case $\varrho \in [1/20, 20]$ and rule QC otherwise.

When computing expressions of the type $\|D_\alpha^k B_\alpha (Au_\alpha - \tilde{f})\|$, the amount of computational work increases proportionally with k. To reduce the amount of computational work, in [29] alternative numerical schemes for computing $\|D_\alpha^k B_\alpha (Au_\alpha - \tilde{f})\|$ are proposed, involving derivatives of u_α

$$\|D_\alpha^k B_\alpha (Au_\alpha - \tilde{f})\| = \begin{cases} \alpha^{-k+1} \dfrac{1}{k!(1+k)^{1/2}} \left\langle \dfrac{\mathrm{d}^k u_\alpha}{\mathrm{d}(\alpha^{-1})^k}, \dfrac{\mathrm{d}^{k+1} u_\alpha}{\mathrm{d}(\alpha^{-1})^{k+1}} \right\rangle^{1/2}, & \text{if } k \in \mathbb{N}, \\ \alpha^{-k+1} \dfrac{1}{(k+1/2)!} \left\| \dfrac{\mathrm{d}^{k+1/2} u_\alpha}{\mathrm{d}(\alpha^{-1})^{k+1/2}} \right\|, & \text{if } k + 1/2 \in \mathbb{N}, \end{cases}$$

their numerical approximation by difference schemes, and linear combinations of approximations u_α with different parameters α.

Let us compare the amount of computation between parameter choice rules using the rules from Figs. 1 and 2. Compared to the discrepancy principle, rule ME needs approximately 2 times as many computations, rules R2 and R(2) 3 times as many and R(8) 9 times as many. In computing our tables we used $\alpha \in [10^{-30}, 1]$. To avoid loss of stability at smaller α's we used alternative computational schemes involving functions Ψ from the Appendix of [24]. In this case ME and R2 still needed 2 and 3 times as many computations than the discrepancy principle but rules R(2) and R(8) needed 5 and 17 times as many computations, respectively. However, for the rules in the tables, the computations needed to use rule R(8) included the computations needed for rules ME, R2, R(2).

6. Conclusion

We propose a general family of rules for a posteriori choice of the regularization parameter, guaranteeing convergence of the regularized approximations as the noise level tends to zero. This family includes all known quasi-optimal rules (monotone error rule, rule R1 (an analog of the balancing principle) etc.) but certain new rules from the family are more stable with respect to the inaccuracies in noise level information than known rules. Numerical comparison of different rules shows, for various inaccuracies of the noise level, which rule should be preferred to other rules. For the case of exact or overestimated noise level we proposed in [22–24] to combine the monotone error rule and rule R2 but the numerical experiments of this paper show that if in this combination a certain new rule from the proposed family is used instead of rule R2, the accuracy of regularized approximation is higher, especially in the case of extensive overestimation of the noise level. For the case of possible extensive under- or overestimation of the noise level in works [23,28,24] we recommended a two-step rule DM, which uses rule R1 and the function from rule R2. It may be possible to improve rule DM using a certain new rule from the general family in one or both steps. If there is no information about the noise level or if the under- or overestimation of the noise level may exceed 20 times, we recommend using the modified quasi-optimality criterion QC. It would be interesting to investigate this family of rules in the case of stochastic noise of the right-hand side.

Acknowledgment

This work was supported by the Estonian Science Foundation, Research Grant No. 7489.

References

[1] H.W. Engl, M. Hanke, A. Neubauer, Regularization of Inverse Problems, in: Mathematics and its Applications, vol. 375, Kluwer Academic Publishers, Dordrecht, 1996.
[2] G.M. Vainikko, A.Yu. Veretennikov, Iteration Procedures in Ill-Posed Problems, Nauka, Moscow, 1986 (in Russian).
[3] A.B. Bakushinskii, Remarks on choosing a regularization parameter using the quasi-optimality and ratio criterion, Comput. Math. Math. Phys. 24 (4) (1984) 181–182.
[4] V.A. Morozov, On the solution of functional equations by the method of regularization, Sov. Math. Dokl. 7 (1966) 414–417.
[5] H. Gfrerer, An a posteriori parameter choice for ordinary and iterated Tikhonov regularization of ill-posed problems leading to optimal convergence rates, Math. Comp. 49 (180) (1987) 507–522.
[6] T. Raus, On the discrepancy principle for solution of ill-posed problems, Acta Comment. Univ. Tartu. 672 (1984) 16–26. (in Russian).
[7] T. Raus, On the discrepancy principle for solution of ill-posed problems with non-selfadjoint operators, Acta Comment. Univ. Tartu. 715 (1985) 12–20. (in Russian).
[8] U. Hämarik, U. Tautenhahn, On the monotone error rule for parameter choice in iterative and continuous regularization methods, BIT 41 (5) (2001) 1029–1038.
[9] U. Tautenhahn, U. Hämarik, The use of monotonicity for choosing the regularization parameter in ill-posed problems, Inverse Problems 15 (6) (1999) 1487–1505.

[10] T. Raus, An a posteriori choice of the regularization parameter in case of approximately given error bound of data, Acta Comment. Univ. Tartu. 913 (1990) 73–87.
[11] T. Raus, About regularization parameter choice in case of approximately given error bounds of data, Acta Comment. Univ. Tartu. 937 (1992) 77–89.
[12] U. Hämarik, T. Raus, About the balancing principle for choice of the regularization parameter, Numer. Funct. Anal. Optim. 30 (9–10) (2009) 951–970.
[13] P. Mathé, The Lepskii principle revisited, Inverse Problems 22 (3) (2006) L11–L15.
[14] P. Mathé, S.V. Pereverzev, Geometry of linear ill-posed problems in variable Hilbert scales, Inverse Problems 19 (3) (2003) 789–803.
[15] U. Hämarik, R. Palm, T. Raus, Use of extrapolation in regularization methods, J. Inverse Ill-Posed Probl. 15 (3) (2007) 277–294.
[16] T. Raus, U. Hämarik, New rule for choice of the regularization parameter in (iterated) Tikhonov method, Math. Model. Anal. 14 (2) (2009) 187–198.
[17] T. Raus, U. Hämarik, On the quasioptimal regularization parameter choices for solving ill-posed problems, J. Inverse Ill-Posed Probl. 15 (4) (2007) 419–439.
[18] F. Bauer, T. Hohage, A Lepskij-type stopping rule for regularized Newton methods, Inverse Problems 21 (6) (2005) 1975–1991.
[19] F. Bauer, S. Kindermann, The quasi-optimality criterion for classical inverse problems, Inverse Problems 24 (035002) (2008).
[20] P. Mathé, S.V. Pereverzev, Regularization of some linear ill-posed problems with discretized random noisy data, Math. Comp. 75 (256) (2006) 1913–1929.
[21] F. Bauer, M.A. Lukas, Comparing parameter choice methods for regularization of ill-posed problems, Math. Comput. Simul. 81 (2011) 1795–1841.
[22] U. Hämarik, R. Palm, T. Raus, On minimization strategies for choice of the regularization parameter in ill-posed problems, Numer. Funct. Anal. Optim. 30 (9–10) (2009) 924–950.
[23] U. Hämarik, R. Palm, T. Raus, Comparison of parameter choices in regularization algorithms in case of different information about noise level, Calcolo 48 (1) (2011) 47–59.
[24] R. Palm, Numerical comparison of regularization algorithms for solving ill-posed problems, Ph.D. Thesis, University of Tartu, 2010. http://hdl.handle.net/10062/14623.
[25] T. Raus, About the general family of the regularization parameter choice rules for Tikhonov and Lavrentiev method, in: Abstracts of the 5th International Conference ''Inverse Problems: Modeling and Simulation'' Held on May 24–29 2010, Antalya, Turkey, Izmir University, 2010, p. 51.
[26] P.C. Hansen, Regularization tools: a Matlab package for analysis and solution of discrete ill-posed problems, Numer. Algorithms 6 (1) (1994) 1–35.
[27] C. Brezinski, G. Rodriguez, S. Seatzu, Error estimates for linear systems with applications to regularization, Numer. Algorithms 49 (1–4) (2008) 85–104.
[28] U. Hämarik, T. Raus, On the choice of the regularization parameter in ill-posed problems with approximately given noise level of data, J. Inverse Ill-Posed Probl. 14 (3) (2006) 251–266.
[29] T. Raus, U. Hämarik, On numerical realization of quasioptimal parameter choices in (iterated) Tikhonov and Lavrentiev regularization, Math. Model. Anal. 14 (1) (2009) 99–108.