A family of rules for the choice of the regularization parameter in the Lavrentiev method in the case of rough estimate of the noise level of the data

Uno Hämarik, Reimo Palm and Toomas Raus

Abstract. We consider the Lavrentiev method for the regularization of linear ill-posed problems with noisy data. Classical rules for the choice of the regularization parameter that use the noise level work well for almost exact noise level information, they fail in the case of the underestimated noise level and give a much larger error than the error by the optimal parameter in the case of the overestimated noise level. Heuristic rules that do not use the noise level give often good results but can not guarantee the convergence. We propose a general family of parameter choice rules, which give good results also in the case of many times under- or overestimated noise level. This family combines heuristic rules and quasi-optimal rules that use the noise level. The quasi-optimality is proved for a sub-family of rules. The advantages of the new rules are demonstrated in extensive numerical experiments.

Keywords. Ill-posed problems, noise level, regularization, Lavrentiev method, regularization parameter choice.

2010 Mathematics Subject Classification. 47A52, 65J20.

1 Introduction

We consider an operator equation

\[ Au = f, \quad f \in R(A), \]

where \( A \in L(H, H) \) is a linear continuous operator, \( A = A^* \geq 0 \), \( H \) is a Hilbert space. In general our problem is ill-posed: the range \( R(A) \) may be non-closed, the kernel \( N(A) \) may be non-trivial. We denote the minimum norm solution by \( u_* \). We suppose that instead of the exact right-hand side \( f \) we have only an approximation \( \tilde{f} \in H \) and some rough estimate \( \delta \) for the noise level \( \| \tilde{f} - f \| \). To compute the regularized solution \( u_\alpha \) of the equation \( Au = f \), we use the Lavrentiev method \([20, 21, 23, 25, 26, 32]\)

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

This work was supported by the Estonian Science Foundation, research grant no. 9120.
In the case of smooth solution \( u_* \) it makes sense to use the iterated Lavrentiev approximation \( u_{m,\alpha} \) for some \( m \in \mathbb{N} \), by iteratively computing solutions of the equations \((\alpha I + A)u_{n,\alpha} = \alpha u_{n-1,\alpha} + \tilde{f}, \ n = 1, 2, \ldots, m; \ u_{0,\alpha} = 0\). The special case \( m = 1 \) gives the Lavrentiev method. Note that problems with non-selfadjoint operator \( A \) are often symmetrized as \( A^* Au = A^* f \) and then regularized by the Lavrentiev method; the corresponding method is the well-known Tikhonov method \( u_\alpha = (\alpha I + A^* A)^{-1} A^* \tilde{f} \).

In the regularization methods an important problem is the choice of a proper regularization parameter \( \alpha \). If \( \alpha \) is too small, the numerical implementation will be unstable and the approximation \( u_\alpha \) will be useless; if \( \alpha \) is too large, then the regularized equation differs from the initial equation too much.

Typically in the theoretical works for the information about the noise level only two cases are considered: 1) the exact noise level \( \delta \) with \( \| \tilde{f} - f \| \leq \delta \) is given; 2) no information about \( \| \tilde{f} - f \| \) is available. We consider the case 3): the approximate noise level \( \delta \) is given, but it is not known whether the inequality \( \| \tilde{f} - f \| \leq \delta \) holds or not. Investigation of this case is motivated by the fact that classical rules (e.g. the modified discrepancy principle \([27]\)) for the choice of the regularization parameter that use the noise level work well for the exact noise level information, but fail in the case of the underestimated noise level and give a much larger error than the error by the optimal parameter in the case of the overestimated noise level.

On the other hand, for heuristic rules \([2–4, 8, 18, 24]\) that do not use the noise level, the convergence of approximate solutions as \( \| \tilde{f} - f \| \rightarrow 0 \) can not be guaranteed (see \([1]\)). To our opinion, the knowledge about the noise level is often between the extreme cases 1) (full information) and 2) (no information) and thus the case 3) needs the attention. We propose for the case 3) rules, which combine in some way the rules that are used in the cases 1) and 2). Namely, our rule will shift the parameter from a quasi-optimal rule, which uses the noise level, to the direction of decrease of the function to be minimized in a certain heuristic rule.

The plan of this paper is the following. In Section 2, we consider some known rules for the a posteriori parameter choice in the Lavrentiev method. In Section 3 we propose a family of rules for the parameter choice; an analogous family of rules for the Tikhonov method was proposed in \([11]\). Section 4 is devoted to the investigation of the behavior of the function used in the family of rules. In Section 5 we show that the new rules from the proposed family combine quaasioptimal rules that use the noise level with heuristic rules that do not use the noise level. In Section 6 we prove for a sub-family of rules the convergence \( \| u_\alpha - u_* \| \rightarrow 0 \) as \( \delta \rightarrow 0 \), provided that \( \| \tilde{f} - f \| / \delta \leq \text{const} \) in the process \( \delta \rightarrow 0 \). Numerical examples in Section 7 show that the rules from the proposed family work well also in the case the noise level is under- or overestimated many times. Note that
extensive numerical comparison of other parameter choice rules can be found in
[3, 8–10, 24].

2 Known rules for a posteriori parameter choice

Many a posteriori parameter choice rules that use the noise level \( \delta, \| \tilde{f} - f \| \leq \delta \)
of the right-hand side can be represented as the following general rule.

**Rule.** Choose the regularization parameter \( \alpha(\delta) \) as the solution of a certain equation

\[
d(\alpha) = b\delta, \quad b > b_0.
\]

If the solution does not exist, then choose \( u_{\alpha(\delta)} = 0 \).

Examples of this general rule are the following rules.

**The modified discrepancy principle (MD rule)** ([27]). Here

\[
d(\alpha) = \| B_\alpha (A u_\alpha - \tilde{f}) \| = \| A u_{2,\alpha} - \tilde{f} \|, \quad b_0 = 1,
\]

where the operator \( B_\alpha \) is defined by \( B_\alpha = \alpha (\alpha I + A)^{-1} \). Note that the discrepancy principle with \( d(\alpha) = \| A u_\alpha - \tilde{f} \| \) and \( b_0 = 1 \) works in the Tikhonov method but fails in the Lavrentiev method.

**The discrepancy of extrapolated approximation (DE rule)** ([6, 7, 9, 26]). Here

\[
d(\alpha) = \| A v_\alpha - \tilde{f} \|, \quad v_\alpha := \frac{u_\alpha - r u_{r \alpha}}{1 - r}
\]

with \( r \neq 1, r > 0 \) is the extrapolated approximation, \( b_0 = 1 \).

**Analogue of the monotone error rule (MEa rule)** ([17, 24]). Here

\[
d(\alpha) = \frac{\| B_\alpha (A u_\alpha - \tilde{f}) \|^2}{\| B_\alpha^2 (A u_\alpha - \tilde{f}) \|} = \frac{\| A u_{2,\alpha} - \tilde{f} \|^2}{\| A u_{3,\alpha} - \tilde{f} \|}, \quad b_0 = 1.
\]

Note that the function \( d(\alpha) \) in these three rules is increasing as \( \alpha \) increases and the equation (2.1) has a unique solution in the case \( \| \tilde{f} \| / b \geq \delta \). In the case of very large noise level \( \delta > \| \tilde{f} \| / b \) the equations in these rules have no solution.

**Another analogue of the monotone error rule (MEd rule)** ([17, 24]). This rule uses the differences \( w_{i,\alpha} := u_{i, r \alpha} - u_{i, \alpha}, r > 1 \), of solutions,

\[
d(\alpha) = \frac{2\alpha}{r - 1} \left( \frac{w_{1,\alpha}, w_{1, r \alpha}}{\| w_{2, r \alpha} \|} \right), \quad n > 1, \quad b_0 = 1.
\]
**Rule R1.** In [28] an a posteriori parameter choice rule was considered for which the operators \( D^k_\alpha \) were applied to the modified discrepancy, where

\[
D_\alpha := \alpha^{-1} A B_\alpha = A (\alpha I + A)^{-1}.
\]

For this rule \( d(\alpha) = \| D^k_\alpha B_\alpha (A u_\alpha - \tilde{f}) \|, \) \( 2k \in \mathbb{N}, k > 0 \) and \( b_0 = \frac{k^k}{(k+2)^{k+2}}. \)

The equation in this rule may have many solutions. Rules ME and DE are quasi-optimal, as well as rule R1 with the smallest solution.

**Theorem 2.1.** Let \( \| \tilde{f} - f \| \leq \delta. \) Let the parameter \( \alpha(\delta) \) be chosen according to rule R1 with smallest solution or rule MD or rule DE. Then for the Lavrentiev method the error estimate

\[
\| u_{\alpha(\delta)} - u_* \| \leq \| u_{\alpha(\delta)}^+ - u_* \| + \frac{\delta}{\alpha(\delta)} \leq C(b) \inf_{\alpha \geq 0} \left( \| u_{\alpha}^+ - u_* \| + \frac{\delta}{\alpha} \right)
\]

holds, where \( u_{\alpha}^+ = (\alpha I + A)^{-1} f. \)

If the rule is quasi-optimal then it is order optimal for any set of the solutions. In many papers for some parameter choice rules the oracle inequality (cf. [2,4,22]) is proved. Note that quasi-optimality of a rule means that for this rule the deterministic oracle inequality holds. The first rule for which the quasi-optimality and oracle inequality was proved is the MD rule (with slightly different coefficients, see [27]) for the Lavrentiev method and some other methods.

Theorem 2.1 is proved for rule DE in [26] and for rules MD, R1 in [30]. For the rule MEa the quasi-optimality is proven in [17] under an additional condition

\[
\| B^2_\alpha (A u_\alpha - \tilde{f}) \| \equiv \| B^3_\alpha A u_* + B^3_\alpha (\tilde{f} - f) \| \geq \| B^3_\alpha (\tilde{f} - f) \| \quad \forall \alpha \geq \alpha(\delta),
\]

where \( \alpha(\delta) \) is the parameter chosen by the MEa rule. This condition is satisfied in special cases if the noise \( \tilde{f} - f \) of the right-hand side is such that \( \tilde{f} - f \in N(A) \) or the scalar product \( (F_\lambda u_*, \tilde{f} - f) \geq 0 \) for all \( \lambda \geq 0, \) where \( \{F_\lambda\} \) is the spectral family of the operator \( A. \) Numerical experiments show that for most severely ill-posed problems in the case of an uniform distribution of the noise \( \tilde{f} - f \) the condition \( \| B^2_\alpha (A u_\alpha - \tilde{f}) \| \geq c_* \| B^3_\alpha (\tilde{f} - f) \| \) is fulfilled with constant \( c_* \approx 1. \)

Rules MD, DE and MEa are unstable with respect to the inaccuracies of the noise level: they fail in the case of underestimated noise level and do not work well in the case of overestimated noise level. Rule R1 allows moderate under- or overestimation of the noise level, but the function \( d(\alpha) \) in these rules is not monotone and the equation (2.1) may have many solutions. In case of exact noise level it is good to use the smallest solution but in case of approximate noise level information the largest solution is better especially in case of underestimated noise level.
Theorem 2.2 ([16, 29]). Let \( \alpha = \alpha(\delta) \) be the largest solution of the equation (2.1), corresponding to rule \( R1 \). If in the process \( \delta \to 0 \) the ratio \( \| \tilde{f} - f \|_\delta \) of the actual and supposed error level is bounded, the convergence is guaranteed:

\[
\| u_{\alpha(\delta)} - u_* \| \leq \| u^+_{\alpha(\delta)} - u_* \| + \frac{\| \tilde{f} - f \|}{\alpha(\delta)} \to 0 \quad \text{as} \quad \delta \to 0.
\]

Note that recommendations for the numerical implementation of the considered and some other rules can be found in [5, 24, 31]. Note also that in the case of the smooth solution \( u_* \) a proper linear combination of \( m \) Lavrentiev approximations (\( v_\alpha \), if \( m = 2 \)) with different values of the parameter \( \alpha \) is a more accurate approximation to \( u_* \) than the single Lavrentiev approximation (see [6, 7, 9, 26]).

3 A family of a posteriori parameter choice rules

In the following we consider a general family of parameter choice rules for the Lavrentiev method, which includes several known quasi-optimal parameter choice rules. Note that an analogous family of parameter choice rules for the Tikhonov method was proposed in [11].

**General rule.** Fix the parameters \( q, k : 4/3 \leq q < \infty, k \geq 0, 2k \in N, 3q \in N \). Let the constant \( b > 1 \) if \( k = 0 \) and \( b > 0 \) if \( k > 0 \). Choose the regularization parameter \( \alpha = \alpha(\delta) \) as the largest solution of the equation

\[
d(\alpha, q, k) := \frac{\kappa_\alpha \| D^{\frac{k}{q}} B^{\frac{k}{q}} (Au_\alpha - \tilde{f}) \|^q / (k-1)}{\| B^{\frac{3q}{2} - 1} (Au_\alpha - \tilde{f}) \|^q / (k-1)} = b \delta,
\]

where

\[
\kappa_\alpha = (1 + \alpha \| A \|^{-1})^{\frac{k q + s_0 q / 2}{q - 1}}, \quad s_0 = \begin{cases} 0, & \text{if } k = 0, \\ 1, & \text{if } k > 0. \end{cases}
\]

If the solution does not exist, then choose \( u_{\alpha(\delta)} = 0 \). Denote this rule by \( R(q, k) \).

The rule (3.1) includes as special cases the MD rule \( (q = 4/3, k = 0) \) and the MEa rule \( (q = 2, k = 0) \).

Note that this 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.
Remark 3.1. The rule (3.1) can be generalized to

\[ d(\alpha, q, k, l) := \frac{\kappa_\alpha \| D^k_\alpha B_\alpha (Au_\alpha - \hat{f}) \|^q/(q-1)}{\| D^3_\alpha B_\alpha^{3q/2-1} (Au_\alpha - \hat{f}) \|^1/(q-1)} = b\delta, \]  

(3.2)

where $4/3 \leq q < \infty$, $l \geq 0$, $k \geq l/q$, $2k \in \mathbb{N}$, $3q \in \mathbb{N}$, $2l \in \mathbb{N}$. This rule includes the rule $R_1$ for $q = 4/3$, $k = l$, the only difference being the multiplier $\kappa_\alpha$; but due to convergence $\kappa_\alpha \to 1$ as $\alpha \to 0$ the difference is marginal at small $\alpha$’s. Note that the factor $\kappa_\alpha$ in the expression of $d(\alpha, q, k)$ improves numerical performance and guarantees the existence of the solution also at larger noise levels $\delta$ (for example $\delta = 0.3\| \hat{f} \|$).

Since the numerical experiments show that the rule $d(\alpha, q, k, 0) = b\delta$ gives better results than the rule $d(\alpha, q, k, l) = b\delta$, $l > 0$ for any $k$ and $q$ (see Table 2 for $q = 4/3$, $k = 1$), we do not investigate the rule (3.2) any further.

The advantage of many rules of the proposed family (3.1) is the stability of the parameter choice with respect to the inaccuracy in the noise level information. This stability depends on the derivative of the function $d(\alpha)$ in the neighborhood of $\alpha(\| \hat{f} - f \|)$, where $\alpha(\| \hat{f} - f \|)$ is the parameter for which $d(\alpha) = b\| \hat{f} - f \|$: the larger the derivative of $d(\alpha)$ is in the neighborhood of $\alpha(\| \hat{f} - f \|)$, the more stable the parameter choice rule $d(\alpha) = b\delta$ is with respect to the inaccuracy in the noise level.

Figure 1 shows the difference in the behavior of functions $d(\alpha) = b^{-1}d(\alpha, q, k)$ with proper $b$ in the example test problem foxgood from [19] (see Section 6 below) in case $\| \hat{f} - f \| = 0.001$. The horizontal axis shows $\alpha$’s, the red line shows the error $\| u_\alpha - u_* \|$ in the Lavrentiev approximation. The parameter $\alpha$ in rule $d(\alpha) = b\delta$ is chosen as the intersection of the corresponding graph of $d(\alpha)/b$ with $\delta$: in case of exact noise level we find the intersection with $\delta = \| \hat{f} - f \|$, in case of 10 times overestimated or 10 times underestimated noise level we take the intersection with $\delta = 10\| \hat{f} - f \|$ or $\delta = 0.1\| \hat{f} - f \|$, respectively. The function $d(\alpha, q, 0)$ in the rule MD is monotone but changes only a little in the neighborhood of $\alpha(\| \hat{f} - f \|)$. Therefore small changes of $\delta$ cause large changes in the corresponding $\alpha(\delta)$’s.

Figure 1 shows that the functions $d(\alpha, 4/3, k)$ with $k > 0$ increase essentially faster than those with $k = 0$ and the corresponding rules $R(4/3, k)$ with $k > 0$ are more stable than the rule MD. In the case of overestimation of the noise level, using, for example, $\delta = 0.01$, the rules $R(4/3, k)$ with $k > 0$ work better than the rule MD. We see that the rules $R(4/3, k)$ have no problem at overestimation of the noise level. Rule MD fails in case of underestimation of the noise level but rules $R(4/3, k)$ with $k > 0$ give good results also at the underestimation: up to $10^3$ times for $k = 1/2$, up to $10^4$ times for $k = 1$, and up to $10^5$ times for $k = 2$. 
A family of rules in the Lavrentiev method

Figure 1. Dependence on \( k \) in the problem \( \text{foxgood}, \| \tilde{f} - f \| = 10^{-3} \).

Note also that for \( \alpha > \alpha(\| \tilde{f} - f \|) \) the function \( d(\alpha, q, k) \) is monotonically increasing in most cases but for \( \alpha \leq \alpha(\| \tilde{f} - f \|) \) the function \( d(\alpha, q, k), k > 0 \), starts to oscillate. Therefore, if the noise level is possibly underestimated, we must take the regularization parameter as the largest solution of the equation (3.1).

4 The behavior of the function \( d(\alpha, q, k) \)

In this section we investigate the function \( d(\alpha) = d(\alpha, q, k) \). We give the limits of this function in the processes \( \alpha \to 0 \) and \( \alpha \to \infty \) and prove the monotonicity of the function for \( k = 0 \).

Lemma 4.1. It holds

\[
d(\alpha, q, k) \leq \kappa_{\alpha} \| D_{\alpha}^{(kq-1)/(q-1)} B_{\alpha}^{(2-q)/(2q-2)} (Au_{\alpha} - \tilde{f}) \|
\]

If \( q \geq 4/3 \), then

\[
d(\alpha, q, 0) \geq \| B_{\alpha} (Au_{\alpha} - \tilde{f}) \|
\]

Proof. To prove the first inequality, we use the equality \( Au_{\alpha} - \tilde{f} = -B_{\alpha} \hat{f} \) and the inequality of moments \( \| Wv \| \leq \| W^{p} v \|^{{1/p}} \| v \|^{1-1/p}, 0 < p \leq 1 \), with the operator \( W = D_{\alpha}^{k} B_{\alpha}^{2-3q/2} \) and the element \( v = B_{\alpha}^{3q/2} \tilde{f} \):

\[
\| D_{\alpha}^{k} B_{\alpha} (Au_{\alpha} - \tilde{f}) \| = \| D_{\alpha}^{k} B_{\alpha}^{2} \hat{f} \|
\leq \| (D_{\alpha}^{k} B_{\alpha}^{2-3q/2} Q)^{q/(q-1)} B_{\alpha}^{3q/2} \hat{f} \|^{1-1/q} \| B_{\alpha}^{3q/2} \hat{f} \|^{1/q}
= \| D_{\alpha}^{kq/(q-1)} B_{\alpha}^{(2-q)/(2q-2)} (Au_{\alpha} - \tilde{f}) \|^{1-1/q} \| B_{\alpha}^{3q/2-1} (Au_{\alpha} - \tilde{f}) \|^{1/q}.
\]
Now we get
\[
d(\alpha, q, k) = \frac{\kappa_\alpha \| D_\alpha^k B_\alpha (A\alpha - \tilde{f}) \|^{q/(q-1)}}{\| B_\alpha^{3q/2-1} (A\alpha - \tilde{f}) \|^{1/(q-1)}}
\leq \kappa_\alpha \| D_\alpha^{kq/(q-1)} B_\alpha^{(2-q)/(2q-2)} (A\alpha - \tilde{f}) \|.
\]

The second inequality follows immediately using the inequalities \( \kappa_\alpha \geq 1 \) and \( \| B_\alpha^{3q/2-1} \| \leq \| B_\alpha \| \) (\( \forall v \in H \)), if \( q \geq 4/3 \):
\[
d(\alpha, q, 0) = \frac{\kappa_\alpha \| B_\alpha (A\alpha - \tilde{f}) \|^{q/(q-1)}}{\| B_\alpha^{3q/2-1} (A\alpha - \tilde{f}) \|^{1/(q-1)}} \geq \| B_\alpha (A\alpha - \tilde{f}) \|.
\]

\[\square\]

**Lemma 4.2.** For the function \( d(\alpha, q, k) \) the following limits hold:

\[
\lim_{\alpha \to 0} d(\alpha, q, k) = 0, \quad \text{if } k > 0, \quad (4.1)
\]
\[
\lim_{\alpha \to 0} d(\alpha, q, 0) \leq \delta, \quad \text{if } q \leq 2 \text{ and } \| \tilde{f} - f \| \leq \delta, \quad (4.2)
\]
\[
\lim_{\alpha \to \infty} d(\alpha, q, k) = \infty, \quad \text{if } k > 0, \quad (4.3)
\]
\[
\lim_{\alpha \to \infty} d(\alpha, q, 0) = \left( \frac{\| A^k \tilde{f} \|^{q}}{\| \tilde{f} \|} \right)^{1/(q-1)}. \quad (4.4)
\]

**Proof.** As \( \| B_\alpha^t \| \leq 1 \) (\( t \geq 0 \)) and \( \lim_{\alpha \to 0} \| D_\alpha^t v \| = 0 \) (\( \forall v \in H, t > 0 \)), we have for \( k > 0 \)
\[
\lim_{\alpha \to 0} \| D_\alpha^{kq/(q-1)} B_\alpha^{(2-q)/(2q-2)} (A\alpha - \tilde{f}) \| = \lim_{\alpha \to 0} \| D_\alpha^{kq/(q-1)} B_\alpha^{1/(2q-2)} \tilde{f} \|
\]
\[
= 0,
\]
\[
\lim_{\alpha \to 0} \| B_\alpha^{(2-q)/(2q-2)} (A\alpha - \tilde{f}) \| \leq \lim_{\alpha \to 0} \| A\alpha - \tilde{f} \| \leq \delta, \quad \text{if } q \leq 2.
\]
Also using the limit \( \lim_{\alpha \to 0} \kappa_\alpha = 1 \), the limits (4.1), (4.2) follow from the results of Lemma 4.1.

The function \( d(\alpha, q, k) \) can be represented in the form
\[
d(\alpha, q, k) = \frac{\kappa_\alpha \alpha^{-kq/(q-1)} \| A^k B_\alpha^{k+1} (A\alpha - \tilde{f}) \|^{q/(q-1)}}{\| B_\alpha^{3q/2-1} (A\alpha - \tilde{f}) \|^{1/(q-1)}}.
\]

As
\[
\lim_{\alpha \to \infty} \alpha^{-kq/(q-1)} \kappa_\alpha
\]
\[
= \lim_{\alpha \to \infty} (\alpha^{-1} + \| A^{-1} \|^{kq/(q-1)} (1 + \alpha \| A \|^{-1})^{kq/(2q-2)}
\]
A family of rules in the Lavrentiev method

\[ \begin{align*}
= \|A\|^{-kq/(q-1)} \lim_{\alpha \to \infty} (1 + \alpha \|A\|^{-1})^{s_0q/(2q-2)} &= \begin{cases} 1, & \text{if } k = 0, \\
\infty, & \text{if } k > 0, 
\end{cases}
\end{align*} \]

and

\[ \lim_{\alpha \to \infty} \|B^t_\alpha (Au_\alpha - \tilde{f})\| = \|\tilde{f}\|, \quad \text{if } t \geq 0, \]

the relations (4.3), (4.4) now follow. \(\square\)

**Lemma 4.3.** The function \(d(\alpha, q, 0), q \geq 4/3, \) as the function of \(\alpha,\) is monotonically increasing.

**Proof.** Showing the monotonicity of the function \(d(\alpha, q, 0)\) is equivalent to showing the monotonicity of \((d(\alpha, q, 0))^{q-1}\). Now for arbitrary \(s \in \mathbb{R}\)

\[ \frac{d}{d\alpha} \|B^s_\alpha \tilde{f}\|^2 = \frac{d}{d\alpha} ((\alpha^{-1} A + I)^{-2s} \tilde{f}, \tilde{f}) \]

\[ = \frac{d}{d\alpha} \int_0^1 \|A\| (\alpha^{-1} \lambda + 1)^{-2s} d \|F_\lambda \tilde{f}\|^2 \]

\[ = \frac{2s}{\alpha^2} \int_0^1 \|A\| \lambda (\alpha^{-1} \lambda + 1)^{-2s-1} d \|F_\lambda \tilde{f}\|^2 \]

\[ = \frac{2s}{\alpha^2} (A(\alpha^{-1} A + I)^{-2s-1} \tilde{f}, \tilde{f}) \]

\[ = \frac{2s}{\alpha^2} \|A^{1/2} B^{s+1/2}_\alpha \tilde{f}\|^2. \]

Using this result for \(s = 2\) and \(s = 3q/2\), respectively, we obtain

\[ \frac{d}{d\alpha} ((d(\alpha, q, 0))^{q-1}) \]

\[ = \frac{d}{d\alpha} \left( \frac{\|B_\alpha (Au_\alpha - \tilde{f})\|^q}{\|B^{3q/2-1}_\alpha (Au_\alpha - \tilde{f})\|^q} \right) = \frac{d}{d\alpha} \left( \frac{\|B^2_\alpha \tilde{f}\|^q}{\|B^{3q/2-2}_\alpha \tilde{f}\|^q} \right) \]

\[ = \frac{2q \|B^2_\alpha \tilde{f}\|^{q-2} \|A^{1/2} B^{5/2}_\alpha \tilde{f}\|^2}{\alpha^2 \|B^{3q/2-2}_\alpha \tilde{f}\|^q} - \frac{(3q/2) \|A^{1/2} B^{3q+1/2}_\alpha \tilde{f}\| \|B^2_\alpha \tilde{f}\|^q}{\alpha^2 \|B^{3q/2-2}_\alpha \tilde{f}\|^3} \]

\[ = \frac{q \|B^2_\alpha \tilde{f}\|^{2-q} \|B^{3q/2-2}_\alpha \tilde{f}\|^{-1}}{2\alpha \|B^{3q/2-2}_\alpha \tilde{f}\|^2} \left( 4\alpha^{-1} \|B^{3/2}_\alpha \tilde{f}\|^2 \|A^{1/2} B^{5/2}_\alpha \tilde{f}\|^2 \right) - 3\alpha^{-1} \|B^2_\alpha \tilde{f}\|^2 \|A^{1/2} B^{3q+1/2}_\alpha \tilde{f}\|^2 \right). \]
Now, by the relation \( \| A^{1/2} B_{\alpha}^{\delta} \tilde{f} \|^2 = \alpha \| B_{\alpha}^{\delta-1/2} \tilde{f} \|^2 - \alpha \| B_{\alpha}^{\delta} \tilde{f} \|^2 \), we get
\[
4 \alpha^{-1} \| B_{\alpha}^{3q/2} \tilde{f} \|^2 \| A^{1/2} B_{\alpha}^{5/2} \tilde{f} \|^2 - 3 \alpha^{-1} \| B_{\alpha}^{\delta} \tilde{f} \|^2 \| A^{1/2} B_{\alpha}^{(3q+1)/2} \tilde{f} \|^2 \\
= 4 \| B_{\alpha}^{3q/2} \tilde{f} \|^2 (\| B_{\alpha}^{\delta} \tilde{f} \|^2 - \| B_{\alpha}^{5/2} \tilde{f} \|^2) \\
- 3 \| B_{\alpha}^{\delta} \tilde{f} \|^2 (\| B_{\alpha}^{3q/2} \tilde{f} \|^2 - \| B_{\alpha}^{(3q+1)/2} \tilde{f} \|^2) \\
= \| B_{\alpha}^{3q/2} \tilde{f} \|^2 (\| B_{\alpha}^{\delta} \tilde{f} \|^2 - \| B_{\alpha}^{5/2} \tilde{f} \|^2) \\
+ 3(\| B_{\alpha}^{3q/2} \tilde{f} \|^2 \| B_{\alpha}^{(3q+1)/2} \tilde{f} \|^2 - \| B_{\alpha}^{3q/2} \tilde{f} \|^2 \| B_{\alpha}^{5/2} \tilde{f} \|^2).
\]

Here, both summands are nonnegative. The nonnegativity of the first term follows from the inequality \( \| B_{\alpha} \| \leq 1 \). The nonnegativity of the second term can be established by multiplying the inequalities
\[
\| B_{\alpha}^{3q/2} \tilde{f} \| \leq \| B_{\alpha}^{(3q+1)/2} \tilde{f} \| \frac{3q-4}{3q-3} \| B_{\alpha}^{5/2} \tilde{f} \| \frac{1}{3q-3}, \\
\| B_{\alpha}^{5/2} \tilde{f} \| \leq \| B_{\alpha}^{(3q+1)/2} \tilde{f} \| \frac{3q-4}{3q-3} \| B_{\alpha}^{3q/2} \tilde{f} \| \frac{1}{3q-3},
\]
which can be obtained from the inequality of moments
\[
\| W^{p} z \| \leq \| W^{q} z \|^{p/r} \| z \|^{1-p/r}, \ W \in L(H, H), \ W = W^* \geq 0, \ z \in H, \ p \leq r
\]
with \( W = B_{\alpha}, \ z = B_{\alpha}^{2} \tilde{f}, \ r = (3q - 3)/2 \) and either \( p = 3q/2 - 2 \) or \( p = 1/2 \). Thus, the function \( d(\alpha, q, 0) \) is monotonically increasing.

These lemmas show that the equation \( d(\alpha, q, 0) = b\delta \) is uniquely solvable, if \( b > b_0 = 1, \delta < (\| A^{k} \tilde{f} \|^{q/\| \tilde{f} \|})^{1/(q-1)} \) and \( \| \tilde{f} - f \| \leq \delta \). If \( k > 0 \), then the equation \( d(\alpha, q, k) = b\delta \) is solvable for every \( b > 0 \), but there may be many solutions.

5 The family of rules combines quasi-optimal and heuristic rules

In the following we show that the rules \( d(\alpha, q, k) = b\delta \) can also be considered as combinations of quasi-optimal rules \( d(\alpha, q, 0) = b\delta, \ b > 1 \) with delta-free rules. A well-known heuristic rule in the Lavrentiev method is the following: take the regularization parameter \( \alpha_{1}(\delta) \) as the local minimizer of the function \( g_{0}(\alpha) = \alpha^{-1} \| B_{\alpha}(Au_{\alpha} - \tilde{f}) \| \). In the non-selfadjoint case in the Tikhonov method the corresponding function \( g_{0}(\alpha) \) has the form \( \alpha^{-1/2} \| B_{\alpha}(Au_{\alpha} - \tilde{f}) \|, \ B_{\alpha} = \alpha^{1/2}(\alpha I + AA^{*})^{-1/2} \), this function is minimized in the Hanke–Raus heuristic rule [18]. In the Tikhonov method good results are also obtained with heuristic rules which minimize the function \( \alpha^{-1/2} \| D_{\alpha}^{k} B_{\alpha}(Au_{\alpha} - \tilde{f}) \|, \ k > 0 \) (if \( k = 1/2 \)
A family of rules in the Lavrentiev method

this is the well-known quasioptimality criterion, cf. [2–4]) but in the Lavrentiev method the minimization of the function

\[ g_k(\alpha) = \alpha^{-1} \| D_\alpha^k B\alpha(Au_\alpha - \tilde{f}) \|, \quad k > 0 \]  

(5.1)
gives not so good results, compared to the minimization of the function \( g_0(\alpha) \). In practice, much better results are obtained by minimizing the function

\[ \tilde{g}_k(\alpha) = \alpha^{-2} \sum_{j=0}^{2k} c_j \| D_\alpha^{j/2} B\alpha(Au_\alpha - \tilde{f}) \|^2, \]
\[ c_j = j/3 + 1, \quad j = 0, 1, \ldots, 2k, \]  

(5.2)
which is the weighted average of the functions \( g_k^2(\alpha) \).

In the following we find the equations whose solutions should be the local minimizers of the functions \( g_k(\alpha) \), \( \tilde{g}_k(\alpha) \).

First let us find the derivative of the function

\[ g_k^2(\alpha) = \alpha^{-2} \| D_\alpha^k B\alpha(Au_\alpha - \tilde{f}) \|^2 = \alpha^2 \| A^k(\alpha I + A)^{-(k+2)} \tilde{f} \|^2. \]

We deduce

\[
\frac{d}{d\alpha} g_k^2(\alpha) \\
= 2\alpha \| A^k(\alpha I + A)^{-(k+2)} \tilde{f} \|^2 - \alpha^2 (2k + 4) \| A^k(\alpha I + A)^{-(k+5/2)} \tilde{f} \|^2 \\
= -2(k + 1)\alpha^{-3} \| A^k(\alpha I + A)^{-(k+2)} \tilde{f} \|^2 \\
+ 2(k + 2)\alpha^{-3} \| A^{k+1/2}(\alpha I + A)^{-(k+5/2)} \tilde{f} \|^2 \\
= 2\alpha^{-3} \left\{ (k + 2) \| D_\alpha^{k+1/2} B\alpha(Au_\alpha - \tilde{f}) \|^2 - (k + 1) \| D_\alpha^k B\alpha(Au_\alpha - \tilde{f}) \|^2 \right\}. \\
(5.3)
\]

Thus the solution of the equation

\[ G_k(\alpha) := \frac{\| D_\alpha^{k+1/2} B\alpha(Au_\alpha - \tilde{f}) \|}{\| D_\alpha^k B\alpha(Au_\alpha - \tilde{f}) \|} = \sqrt{\frac{k + 1}{k + 2}} \]

is a local minimum of the function \( g_k(\alpha) \) if the function \( G_k(\alpha) \) is increasing in the neighborhood of the solution.

Using (5.3) it is easy to show that the derivative of the function \( \tilde{g}_k(\alpha) \) with weights \( c_j = j/3 + 1, \quad j = 0, 1, \ldots, 2k \) is
\[
\frac{d}{d\alpha} \tilde{g}_k(\alpha) = 2\alpha^{-3} \left\{ \varepsilon_k \| D_\alpha^{k+1/2} B_\alpha (A u_\alpha - \tilde{f}) \|^2 - \| B_\alpha (A u_\alpha - \tilde{f}) \|^2 \right\}, \]
\[
\varepsilon_k = c_{2k} (k + 2) = (3 + 2k)(k + 2)/3
\]
and thus the solution of the equation
\[
\tilde{G}_k(\alpha) := \frac{\| D_\alpha^{k+1/2} B_\alpha (A u_\alpha - \tilde{f}) \|}{\| B_\alpha (A u_\alpha - \tilde{f}) \|} = \varepsilon_k^{-1/2}
\]
is a local minimum of the function \( \tilde{g}_k(\alpha) \) if the function \( \tilde{G}_k(\alpha) \) is increasing in some neighborhood of the solution.

The function \( d(\alpha, q, k) \) can be represented in the form
\[
d(\alpha, q, k) = \frac{\kappa_\alpha \| D_\alpha^k B_\alpha (A u_\alpha - \tilde{f}) \|_{1/(q-1)}}{\| B_\alpha^{3q/2-1} (A u_\alpha - \tilde{f}) \|_{1/(q-1)}}
\]
\[
= \frac{\kappa_\alpha \| B_\alpha (A u_\alpha - \tilde{f}) \|_{q/(q-1)} \left( \| D_\alpha^k B_\alpha (A u_\alpha - \tilde{f}) \| \right)_{q/(q-1)}}{\| B_\alpha^{3q/2-1} (A u_\alpha - \tilde{f}) \|_{1/(q-1)}}
\]
\[
= \kappa_\alpha d(\alpha, q, 0)(\tilde{G}_k_{1/2}(\alpha))^{q/(q-1)}.
\]

Thus the rule \( d(\alpha, q, k) = b\delta \) can be obtained by multiplying the function \( d(\alpha, q, 0) \) of a quasi-optimal rule \( d(\alpha, q, 0) = b\delta \) with the adjusting factor \( (\tilde{G}_k_{1/2}(\alpha))^{q/(q-1)} \) that corrects the parameter choice rule. Namely, if a quasi-optimal rule \( d(\alpha, q, 0) = b\delta \) gives the parameter \( \alpha(\delta) \) but the function \( \tilde{g}_{k-1/2}(\alpha) \) of the delta-free rule is still monotonically increasing at \( \alpha(\delta) \) (i.e. \( \tilde{G}_{k-1/2}(\alpha) > \varepsilon_{k-1/2}^{-1/2} \)), then the rule
\[
d(\alpha, q, k) = \varepsilon_{k-1/2}^{-q/(2q-2)} b\delta
\]
chooses a parameter smaller than \( \alpha(\delta) \); if the function \( \tilde{g}_{k-1/2}(\alpha) \) of the rule is monotonically decreasing at \( \alpha(\delta) \), then the chosen parameter is larger than \( \alpha(\delta) \). Since in the rule \( d(\alpha, q, 0) = b\delta \) the coefficient \( b \) must satisfy the inequality \( b > 1 \), in the rule \( d(\alpha, q, k) = b\delta, k > 0 \) the parameter \( b \) should satisfy \( b > \varepsilon_{k-1/2}^{-q/(2q-2)} \). But numerical experiments show that better results are obtained by taking \( b \) larger:
\[
b \approx \sqrt{k(2q - 1)} \varepsilon_{k-1/2}^{-q/(2q-2)}.
\]

Since at smaller \( q \) the exponent of the stabilizing function \( \tilde{G}_{k-1/2}(\alpha) \) is larger, the rule with \( q = 4/3 \) is most stable with respect to inaccuracy of the noise level of the initial data. Also Figure 2 shows that the rule with \( q = 4/3 \) is more stable than the rules with larger \( q \). In the case of exact error bound \( \delta = \| \tilde{f} - f \| \) the accuracy
A family of rules in the Lavrentiev method

843

Figure 2. Dependence on $q$ in the problem $\text{foxgood}$, $\| \tilde{f} - f \| = 10^{-3}$.

of rules $d(\alpha, q, k) = b\delta$ somewhat increases with increasing $q$ but improvement in accuracy is marginal for $q > 7/3$. Taking into account both stability and accuracy properties, we recommend to use the values $q = 4/3, q = 5/3, q = 2$ or $q = 7/3$; the less information there is about the noise level, the smaller should be $q \in [4/3, 7/3]$.

Note that in our earlier works [12–15, 24] we combined quasi-optimal and heuristic rules in a two-step algorithm DM, where on the first step the parameter $\alpha_{R1}$ is computed by the rule R1 and on the second step the regularization parameter is found as the minimizer of a certain function on the interval $[\alpha_{R1}, 1]$.

6 Convergence and quasi-optimality of family of rules

In this section we show that the choice of the regularization parameter by the proposed family of rules for the Lavrentiev method guarantees the convergence, stability in the case $k > 0$ and the quasi-optimality in the case $k = 0$.

Theorem 6.1. Let $\alpha = \alpha(\delta)$ be the solution of the equation $d(\alpha, q, k) = b\delta$,

$$b > b_0 := \frac{(2k)^kq/(q-1)}{(2k+1)(2k+1)q/(2q-2)}.$$

$k \geq 0, q < 2$. If $\| \tilde{f} - f \| \leq \delta$, then

$$\| u_{\alpha(\delta)} - u_* \| \to 0 \quad \text{as } \delta \to 0.$$
Proof. At first we give some auxiliary results:

\[ \| u_\alpha - u_* \| \leq \| u_\alpha^+ - u_* \| + \frac{\delta}{\alpha}, \]  
\[ Au_\alpha - \tilde{f} = -B_\alpha Au_* - B_\alpha (\tilde{f} - f), \]
\[ \| D_\alpha^k B_\alpha^2 Au_* \| \leq \| D_\alpha^k B_\alpha (Au_\alpha - \tilde{f}) \| + \| D_\alpha^k B_\alpha^2 (\tilde{f} - f) \| \]
\[ \leq \| D_\alpha^k B_\alpha (Au_\alpha - \tilde{f}) \| + \delta. \]  
(6.1)

Consider the process \( \delta \to 0 \). From the equality \( d(\alpha(\delta), q, k) = b \delta \) and from the inequalities \( \kappa_\alpha \geq 1, \| B_\alpha^{3q/2-1} (Au_\alpha - \tilde{f}) \| \leq \| \tilde{f} \| \) it follows
\[ \| D_\alpha^k B_\alpha(\delta) (Au_\alpha(\delta) - \tilde{f}) \| \to 0, \]
therefore, by using the inequality (6.2), we get
\[ \| D_\alpha^k B_\alpha^2(\delta) Au_* \| \to 0 \]  
as \( \delta \to 0 \).

If in the process \( \| D_\alpha^k B_\alpha^2(\delta) Au_* \| \to 0 \) the parameter \( \alpha(\delta) \to 0 \), then
\[ \| u_\alpha^+ - u_* \| = \| B_\alpha(\delta) u_* \| \to 0. \]

Consider now the case where in the process \( \| D_\alpha^k B_\alpha^2(\delta) Au_* \| \to 0 \) the parameter \( \alpha(\delta) \geq \bar{\alpha} = \text{const} > 0 \). In [32, Lemma 3.2, p. 66] it is proved that if \( AB_{\alpha_n}u_* \to 0 \) \( (n \to \infty), \alpha_n \geq \bar{\alpha} = \text{const} > 0 \), then \( B_{\alpha_n}u_* \to 0 \) \( (n \to \infty) \). Similarly we can show that if
\[ D_{\alpha_n}^k B_{\alpha_n}^2 Au_* = \bar{\alpha}_n^{-k} A_{\alpha_n}^{k+1} B_{\alpha_n}^{k+2} u_* \to 0 \]  
\( (n \to \infty), \)
then \( B_{\alpha_n}u_* \to 0 \) \( (n \to \infty) \). Hence, \( \| u_\alpha^+ - u_* \| \to 0 \) \( (\delta \to 0) \) and the convergence of the first term of (6.1) is proved. To prove the convergence of the second term of (6.1), we use Lemma 4.1.

Let the regularization parameter \( \alpha_1(\delta) \) be the minimal solution of the equation
\[ \| D_\alpha^{kq/(q-1)} B_\alpha^{(2-q)/(2q-2)} (Au_\alpha - \tilde{f}) \| = c \delta, \quad b_0 < c < b. \]
If \( k \geq 0 \), \( \| \tilde{f} - f \| \leq \delta \), then
\[ \| D_\alpha^{kq/(q-1)} B_\alpha^{q/(2q-2)} (\tilde{f} - f) \| \leq b_0 \delta \]
and
\[ \| B_\alpha^{(2-q)/(2q-2)} Au_* \| \geq \| D_\alpha^{kq/(q-1)} B_\alpha^{(2-q)/(2q-2)} B_\alpha Au_* \| \]
\[ \geq (c - b_0) \delta. \]  
(6.3)
If \( q < 2 \), then (see [32, p. 43, Lemma 5.2])

\[
\alpha^{-1} \| B_{\alpha}^{(2-q)/(2q-2)} B_{\alpha} A u_* \| \to 0, \quad \text{if } \alpha \to 0.
\] (6.4)

Now from relations (6.3), (6.4) we get \( \delta/\alpha_1(\delta) \to 0 \), if \( \alpha_1(\delta) \to 0 \) in the process \( \delta \to 0 \). If \( \alpha_1(\delta) \geq \bar{\alpha} > 0 \) in the process \( \delta \to 0 \), then the convergence \( \delta/\alpha_1(\delta) \to 0 \) is obvious.

By using Lemma 4.1, we have for sufficiently small \( \delta \) for every \( \alpha \leq \alpha_1(\delta) \)

\[
d(\alpha, q, k) \leq \kappa_{\alpha} \| D_{\alpha}^{kq/(q-1)} B_{\alpha}^{(2-q)/(2q-2)} (A u_{\alpha} - \tilde{f}) \| \leq \kappa_{\alpha} c \delta \leq b\delta.
\]

As \( \alpha(\delta) \) is the solution of the equation \( d(\alpha, q, k) = b\delta \), it follows from the last inequality that \( \alpha(\delta) \geq \alpha_1(\delta) \), which with the convergence \( \delta/\alpha_1(\delta) \to 0 (\delta \to 0) \) proves the convergence of the second term of (6.1).

In the previous theorem we showed that for any solution \( \alpha(\delta) \) of the equation \( d(\alpha, q, k) = b\delta \) the convergence \( \| u_{\alpha(\delta)} - u_* \| \to 0 \) holds, if \( b > b_0 \) and \( \| \tilde{f} - f \| \leq \delta \). In the following theorem we prove that if \( k > 0 \) and we choose the regularization parameter as the largest solution, then the convergence holds also in the case of underestimated noise level.

**Theorem 6.2.** Let \( \alpha = \alpha(\delta) \) be the largest solution of \( d(\alpha, q, k) = b\delta \), \( b = \text{const}, k > 0, q < 2 \). If \( \| \tilde{f} - f \| \leq c_* = \text{const} \) in the process \( \delta \to 0 \), then

\[
\| u_{\alpha(\delta)} - u_* \| \to 0 \quad (\delta \to 0).
\]

**Proof.** The proof of the convergence of the first term of (6.1) is the same as in Theorem 6.1. In the following we consider the convergence of the second term of (6.1).

In [28] it is proved that if we choose the parameter \( \alpha_1(\delta) \) as the largest solution of the equation \( \| D_{\alpha}^k B_{\alpha} (A u_{\alpha} - \tilde{f}) \| = c \delta, k > 0, c > 0 \), then in case \( \| \tilde{f} - f \| \leq c_* \) the convergence \( \delta/\alpha_1(\delta) \to 0 (\delta \to 0) \) is valid. Analogously, we can show the convergence

\[
\delta/\alpha_1(\delta) \to 0 \quad (\delta \to 0)
\] (6.5)

in the case we choose the parameter \( \alpha_1(\delta) \) as the largest solution of the equation

\[
\| D_{\alpha}^k B_{\alpha} (A u_{\alpha} - \tilde{f}) \| = c \delta, k > 0, c > 0, t > 0.
\]

Taking \( c = b - \epsilon, \epsilon > 0 \) and using Lemma 4.1, we get for sufficiently small \( \delta \)

\[
d(\alpha_1(\delta), q, k) \leq \kappa_{\alpha_1(\delta)} \| D_{\alpha_1(\delta)}^{kq/(q-1)} B_{\alpha_1(\delta)}^{(2-q)/(2q-2)} (A u_{\alpha_1(\delta)} - \tilde{f}) \|
\]

\[
= \kappa_{\alpha_1(\delta)} (b - \epsilon) \delta \leq b\delta.
\]
As \( \lim_{\alpha \to \infty} d(\alpha, q, k) = \infty \), if \( k > 0 \), it follows from the last inequality that \( \alpha(\delta) \geq \alpha_1(\delta) \), which together with (6.5) proves the convergence of the second term of (6.1).

In the following we need the condition

\[
\| B_{\alpha}^{3q/2-1} (Au_{\alpha} - \tilde{f}) \| \geq \| B_{\alpha}^{3q/2}(\tilde{f} - f) \| \quad \forall \alpha \geq \alpha(\delta),
\]

(6.6)

where \( \alpha(\delta) \) is the largest solution of the equation \( d(\alpha, q, k) = b\delta \). This condition is a generalization of the condition (2.3), where \( q = 2 \) was used.

**Theorem 6.3.** Let \( \| \tilde{f} - f \| \leq \delta \) and \( b > (\sqrt{2k\alpha})^{q/(q-1)}b_0 \). Let the assumption (6.6) be satisfied. If

\[
d(\alpha, q, k) \geq b\delta \quad \forall \alpha \geq \alpha_0,
\]

then the function

\[
t_c(\alpha) = \| u_{\alpha}^+ - u_* \|^2 + c \| u_{\alpha}^+ - u_{\alpha} \|^2,
\]

c = 0.5b^{(2q-2)/q} - \beta_k^2 \gamma \alpha, \beta_k := b_0^{(q-1)/q}, \gamma \alpha := \kappa_{\alpha}^{(2q-2)/q}, \text{ is monotonically increasing for } \alpha \geq \alpha_0.

**Proof.** First we show that if

\[
\tau_{\alpha}^{k+1/2} \| D_{\alpha}^kB_{\alpha}^2Au_* \| \geq \sqrt{c} \| B_{\alpha}^{3/2}(\tilde{f} - f) \|,
\]

c > 0, for all \( \alpha \in [\alpha_0, \alpha_1] \), \( \tau_{\alpha} := 1 + \alpha \| A \|^{-1} \), then the function

\[
t_c(\alpha) = \| u_{\alpha}^+ - u_* \|^2 + c \| u_{\alpha}^+ - u_{\alpha} \|^2 = \| B_{\alpha}u_* \|^2 + c \alpha^{-2} \| B_{\alpha}(\tilde{f} - f) \|^2
\]

is monotonically increasing on the interval \([\alpha_0, \alpha_1]\). As \( \frac{d}{d\alpha} B_{\alpha} = \alpha^{-2}AB_{\alpha}^2 \), we have

\[
t_c(\alpha) = 2\alpha^{-2} \| B_{\alpha}^{3/2}A^{1/2}u_* \|^2 - 2c\alpha^{-3} \| B_{\alpha}^{3/2}(\tilde{f} - f) \|^2.
\]

As \( \alpha^{-1} \| AB_{\alpha} \| \leq \| A \|/(\alpha + \| A \|) = (1 + \alpha \| A \|^{-1})^{-1} = \tau_{\alpha}^{-1} \leq 1 \), then

\[
t_c'(\alpha) \geq 2\alpha^{-2} \tau_{\alpha}^{2k+1} \| \alpha^{-1} \| B_{\alpha}^{k+1/2}A^{1/2}B_{\alpha}^{3/2}A^{1/2}u_* \|^2
\]

\[
- 2c\alpha^{-3} \| B_{\alpha}^{3/2}(\tilde{f} - f) \|^2
\]

\[
= 2\alpha^{-3} \tau_{\alpha}^{2k+1} \| D_{\alpha}^kB_{\alpha}^2Au_* \|^2 - 2c\alpha^{-3} \| B_{\alpha}^{3/2}(\tilde{f} - f) \|^2,
\]

from which the monotonic increase of the function follows.
A family of rules in the Lavrentiev method

The following inequalities hold:

\[ \| D^k_{\alpha} B_{\alpha} (Au - \tilde{f}) \|_2 \leq 2(\| D^k_{\alpha} B_{\alpha}^2 Au_* \|^2 + \| D^k_{\alpha} B_{\alpha}^2 (\tilde{f} - f) \|^2) , \]

\[ \| D^k_{\alpha} B_{\alpha}^2 (\tilde{f} - f) \| \leq \| D^k_{\alpha} B_{\alpha}^{1/2} \| B_{\alpha}^{3/2} (\tilde{f} - f) \| , \]

\[ \| B_{\alpha}^{3/2} (\tilde{f} - f) \| \leq \| B_{\alpha}^{3q/2} (\tilde{f} - f) \|^{1/q} \| \tilde{f} - f \|^{1-1/q} \]

\[ \leq \| B_{\alpha}^{3q/2} (\tilde{f} - f) \|^{1/q} \delta^{1-1/q} . \]

Because

\[ \| D^k_{\alpha} B_{\alpha}^{1/2} \| = \alpha^{1/2} \max_{0 \leq \lambda \leq \| A \|} \frac{\lambda^k}{(\alpha + \lambda)^{k+1/2}} = \beta_k , \]

we get, by taking into account the above inequalities and condition (6.6), that

\[ d(\alpha, q, k)^{(2q-2)/q} = \frac{\gamma_{\alpha} \| D^k_{\alpha} B_{\alpha} (Au - \tilde{f}) \|^2}{\| B_{\alpha}^{3q/2-1} (Au - \tilde{f}) \|^2/q} \]

\[ \leq \frac{2\gamma_{\alpha} (\| D^k_{\alpha} B_{\alpha}^2 Au_* \|^2 + \| D^k_{\alpha} B_{\alpha}^2 (\tilde{f} - f) \|^2)}{\| B_{\alpha}^{3q/2} (\tilde{f} - f) \|^2/q} \]

\[ \leq \frac{2\gamma_{\alpha} (\| D^k_{\alpha} B_{\alpha}^2 Au_* \|^2 + \beta_k^2 \| B_{\alpha}^{3q/2} (\tilde{f} - f) \|^{2/q} \delta^{(2q-2)/q})}{\| B_{\alpha}^{3q/2} (\tilde{f} - f) \|^2/q} \]

\[ \leq \frac{2\gamma_{\alpha} \| D^k_{\alpha} B_{\alpha}^2 Au_* \|^2}{\| B_{\alpha}^{3q/2} (\tilde{f} - f) \|^2/q} + 2\beta_k^2 \gamma_{\alpha} \delta^{(2q-2)/q} , \]

which by using the inequality \( d(\alpha, q, k) \geq b \delta \) gives

\[ \gamma_{\alpha} \| D^k_{\alpha} B_{\alpha}^2 Au_* \|^2 \geq (b^{(2q-2)/q} / 2 - \beta_k^2 \gamma_{\alpha}) \delta^{(2q-2)/q} \| B_{\alpha}^{3q/2} (\tilde{f} - f) \|^{2/q} \]

\[ \geq (b^{2q-2}/2 - \beta_k^2 \gamma_{\alpha}) \| B_{\alpha}^{3q/2} (\tilde{f} - f) \|^2 . \]

Since \( \gamma_{\alpha}^{1/2} = \tau_{\alpha}^{k+s_0/2} \leq \tau_{\alpha}^{k+1/2} \), the last inequalities imply

\[ \tau_{\alpha}^{k+1/2} \| D^k_{\alpha} B_{\alpha}^2 Au_* \| \geq (b^{(2q-2)/q} / 2 - \beta_k^2 \gamma_{\alpha})^{1/2} \| B_{\alpha}^{3q/2} (\tilde{f} - f) \| , \]

so the function \( t_c(\alpha) , c = 0.5b^{2q-2}/q - \beta_k^2 \gamma_{\alpha} \) is monotonically increasing for all \( \alpha \geq \alpha_0 \).

From Theorem 6.3 we conclude that if \( \| \tilde{f} - f \| \leq \delta \) and the coefficient \( b \) is sufficiently large, then choosing the largest solution of the equation \( d(\alpha, q, k) = b \delta \) as the regularization parameter, we do not take a too small regularization parameter \( \alpha \) at which the function \( t_c(\alpha) \) already starts to decrease monotonically.
Theorem 6.3 also implies that if we choose the regularization parameter $\alpha(\delta)$ as the largest solution of the equation $d(\alpha, q, k) = b\delta$, $b > (\sqrt{2\kappa_\alpha})^{q/q-1}b_0$, then in the case $\|\tilde{f} - f\| \leq \delta$ the convergence $\|u_{\alpha(\delta)} - u_*\| \to 0$ also holds if $q \geq 2$.

The following theorem proves the quasi-optimality of rules $d(\alpha, q, 0) = b\delta$ for any $q > 4/3$. Note that for $q = 4/3$ and $q = 2$ the quasi-optimality has been proved in [17, 30].

**Theorem 6.4.** Let $\alpha = \alpha(\delta)$ be the solution of the equation $d(\alpha, q, 0) = b\delta$, $b > 2^{q/(2q-2)}$, $q > 4/3$. Let the assumption (2.3) be satisfied. Then

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

**Proof.** We use the inequality

$$
\|u_\alpha - u_*\| \leq \|u_\alpha^+ - u_*\| + \|u_\alpha^- - u_\alpha\|
$$

and estimate both terms on the right-hand side separately.

1. Let the parameter $\alpha_1(\delta)$ be chosen by the MD rule

$$
d(\alpha, 4/3, 0) = \|B_\alpha(Au_\alpha - \tilde{f})\| = b\delta.
$$

Since the function $d(\alpha) = d(\alpha, q, 0)$ is monotonically increasing, we have by Lemma 4.1

$$
d(\alpha, q, 0) \geq \|B_\alpha(Au_\alpha - \tilde{f})\| \geq b\delta, \quad \alpha \geq \alpha_1(\delta).
$$

Thus $\alpha(\delta) \leq \alpha_1(\delta)$ and

$$
\|u_{\alpha(\delta)}^+ - u_*\| \leq \|u_{\alpha_1(\delta)} - u_*\| \leq C_{\text{MD}}(b) \inf_{\alpha \geq 0} \left\{ \|u_\alpha^+ - u_*\| + \frac{\delta}{\alpha} \right\},
$$

since the MD rule is quasi-optimal.

2. To estimate the second summand in (6.7), we note that if $\alpha \geq \alpha(\delta)$, then by Theorem 6.3 the function $t_c(\alpha)$, $c := C_1(b) = 0.5b(2q-2)/q - 1$, is monotonically increasing at $\alpha \geq \alpha(\delta)$. Let $\alpha_*$ be the global minimizer of $t_c(\alpha)$. As the function $t_1(\alpha) = \|u_\alpha^+ - u_\alpha\|$ is monotonically decreasing, we get

$$
\|u_{\alpha(\delta)}^+ - u_{\alpha(\delta)}^-\| \leq \|u_{\alpha_*}^+ - u_{\alpha_*}\|
\leq \max(c^{-1}, 1)t_c(\alpha_*).
\leq \max(c^{-2}, 1) \inf_{\alpha \geq 0} \left\{ \|u_\alpha^+ - u_*\| + \|u_\alpha^- - u_\alpha\| \right\}.
$$

Now using the inequalities (6.7)–(6.9), we conclude that the claim of the theorem holds with $C(b) = C_{\text{MD}}(b) + \max(C_{1}^{-2}(b), 1)$. \(\square\)
7 Numerical results

Our tests are performed on the well-known set of test problems by Hansen [19]: \textit{baart, deriv2, foxgood, gravity, heat, ilaplace, phillips, shaw, spikes, wing}. After discretization (discretization parameter \( N = 100 \)) all problems were scaled (normalized) in such a way that the Euclidian norms of the operator and the right-hand side were 1. On base of exact data \( f \) we formed the noisy data \( \tilde{f} \), where \( \| \tilde{f} - f \| \) had values \( 0.3, 10^{-1}, 10^{-2}, \ldots, 10^{-7} \). The noise \( \tilde{f} - f \) added to \( f \) had normal distribution. Here, \( \| \cdot \| \) means the Euclidean norm.

We generated 10 noise vectors and used these vectors in all problems. The problems were regularized by the Lavrentiev method, choosing the regularization parameters by the family of rules using rough estimate \( \delta \) of the noise level.

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 the rule \( R \) on this particular problem. To compare the rules or to present their properties, we present the averages of these error ratios over various parameters of the data set (problems 1–10, noise levels \( \delta \), runs). Tables 1, 3 and Figure 3 give results for rule \( R(q, k) \), corresponding to the functions \( d(\alpha, q, k) \), Table 2 corresponds to the rule (3.2). In all tables the header contains values of \( \rho = \delta / \| \tilde{f} - f \| \). Table 1 shows that in rule \( R(q, 0) \) the accuracy increases somewhat for increasing \( q \) on the interval \([4/3, 3]\). Table 2 shows that the results at \( l > 0 \) are worse than in the case \( l = 0 \). Among the rules \( R(q, 1) \), \( q \in [4/3, 7/3] \), the rule \( R(4/3, 1) \) is most stable and the rule \( R(7/3, 1) \) is most accurate. At underestimation up to 16 times the rule \( R(5/3, 1/2) \) gives normal results but it is not so good at underestimation up to 64 times. Recall that the rule \( R(q, 0) \) fails at underestimated noise level.

In Figure 3 the horizontal axis displays values of the accuracy \( \rho = \delta / \| \tilde{f} - f \| \) of the noise level in the interval \( \rho \in [2^{-6}, 2^7] \), the graphs show error ratios in

<table>
<thead>
<tr>
<th>( q \setminus \rho )</th>
<th>1</th>
<th>( 2^{0.25} )</th>
<th>( 2^{0.5} )</th>
<th>( 2^{0.75} )</th>
<th>( 2^{1} )</th>
<th>( 2^{1.25} )</th>
<th>( 2^{1.5} )</th>
<th>( 2^{1.75} )</th>
<th>( 2^{2} )</th>
<th>( 2^{2.25} )</th>
<th>( 2^{2.5} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>4/3</td>
<td>1.31</td>
<td>1.33</td>
<td>1.39</td>
<td>1.45</td>
<td>1.55</td>
<td>1.64</td>
<td>1.73</td>
<td>1.83</td>
<td>1.92</td>
<td>2.06</td>
<td>2.18</td>
</tr>
<tr>
<td>5/3</td>
<td>1.11</td>
<td>1.11</td>
<td>1.14</td>
<td>1.18</td>
<td>1.24</td>
<td>1.29</td>
<td>1.35</td>
<td>1.40</td>
<td>1.47</td>
<td>1.56</td>
<td>1.64</td>
</tr>
<tr>
<td>2</td>
<td>1.05</td>
<td>1.05</td>
<td>1.09</td>
<td>1.13</td>
<td>1.18</td>
<td>1.24</td>
<td>1.29</td>
<td>1.34</td>
<td>1.40</td>
<td>1.47</td>
<td>1.55</td>
</tr>
<tr>
<td>7/3</td>
<td>1.04</td>
<td>1.05</td>
<td>1.08</td>
<td>1.11</td>
<td>1.16</td>
<td>1.22</td>
<td>1.26</td>
<td>1.31</td>
<td>1.37</td>
<td>1.44</td>
<td>1.52</td>
</tr>
<tr>
<td>3</td>
<td>1.03</td>
<td>1.04</td>
<td>1.07</td>
<td>1.11</td>
<td>1.16</td>
<td>1.21</td>
<td>1.26</td>
<td>1.31</td>
<td>1.37</td>
<td>1.43</td>
<td>1.51</td>
</tr>
</tbody>
</table>

Table 1. Error ratios in rule \( R(q, 0) \) at \( \rho \geq 1 \).
Table 2. Dependence of error ratios on \( l \) in rule (3.2) with \( q = 4/3, k = 1 \).

<table>
<thead>
<tr>
<th>( l \backslash q )</th>
<th>1/64</th>
<th>1/16</th>
<th>1/4</th>
<th>1/2</th>
<th>1</th>
<th>2</th>
<th>4</th>
<th>16</th>
<th>64</th>
<th>256</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>2.19</td>
<td>1.29</td>
<td>1.13</td>
<td>1.10</td>
<td>1.09</td>
<td>1.11</td>
<td>1.17</td>
<td>1.45</td>
<td>2.15</td>
<td>3.76</td>
</tr>
<tr>
<td>1/2</td>
<td>22.3</td>
<td>2.55</td>
<td>1.57</td>
<td>1.35</td>
<td>1.26</td>
<td>1.34</td>
<td>1.53</td>
<td>2.36</td>
<td>3.93</td>
<td>5.92</td>
</tr>
<tr>
<td>1</td>
<td>1e3</td>
<td>2.29</td>
<td>1.28</td>
<td>1.28</td>
<td>1.42</td>
<td>1.73</td>
<td>2.11</td>
<td>3.43</td>
<td>5.29</td>
<td>7.77</td>
</tr>
</tbody>
</table>

Table 3. Error ratios in rules \( R(q, k) \).

<table>
<thead>
<tr>
<th>( q \backslash k )</th>
<th>1/64</th>
<th>1/16</th>
<th>1/4</th>
<th>1/2</th>
<th>1</th>
<th>2</th>
<th>4</th>
<th>16</th>
<th>64</th>
<th>256</th>
</tr>
</thead>
<tbody>
<tr>
<td>4/3</td>
<td>2.81</td>
<td>1.35</td>
<td>1.17</td>
<td>1.11</td>
<td>1.09</td>
<td>1.10</td>
<td>1.13</td>
<td>1.35</td>
<td>1.93</td>
<td>3.30</td>
</tr>
<tr>
<td>5/3</td>
<td>58.0</td>
<td>1.56</td>
<td>1.05</td>
<td>1.04</td>
<td>1.07</td>
<td>1.17</td>
<td>1.34</td>
<td>1.98</td>
<td>3.46</td>
<td>5.34</td>
</tr>
<tr>
<td>5/3</td>
<td>5</td>
<td>1.30</td>
<td>1.08</td>
<td>1.04</td>
<td>1.05</td>
<td>1.10</td>
<td>1.21</td>
<td>1.69</td>
<td>2.96</td>
<td>4.81</td>
</tr>
<tr>
<td>2</td>
<td>8.14</td>
<td>1.38</td>
<td>1.08</td>
<td>1.03</td>
<td>1.04</td>
<td>1.10</td>
<td>1.22</td>
<td>1.77</td>
<td>3.22</td>
<td>5.09</td>
</tr>
<tr>
<td>7/3</td>
<td>2e2</td>
<td>1.99</td>
<td>1.15</td>
<td>1.05</td>
<td>1.02</td>
<td>1.05</td>
<td>1.15</td>
<td>1.60</td>
<td>2.78</td>
<td>4.80</td>
</tr>
</tbody>
</table>

rules \( R(4/3, 0) \), \( R(4/3, 1) \) and \( R(2, 1) \). Since the performance of methods and rules generally depends on the smoothness \( p \) of the exact solution, we also show error ratios for smoothened solutions \( A^pu_\ast (p = 2) \) (computing the right-hand side as \( A(A^pu_\ast) \)) by the dashed lines in this figure. This figure shows that the best constants \( b \) in the case \( p = 2 \) are somewhat larger than those in the case \( p = 0 \). The constants we selected are a compromise for both cases. Figure 3 does not show the graph of the rule \( R(4/3, 0) \) (the MD rule) at \( \rho < 1 \) due to failure of MD. Rules \( R(4/3, 1) \) and \( R(2, 1) \) give much better results than the rule MD in case of exact or overestimated noise level and good results also in the case of underestimated noise level.

Note that rule ME\( d \) with \( r = 1.1, n = 6, b_0 = 1.15 \) gave the error ratio 1.02 for exact data (for both \( p = 0 \) and \( p = 2 \)) and the error ratio 1.45 for 8 times overestimated noise level.

Table 4 compares results of heuristic rules with minimizing functions (5.1) and (5.2), respectively. We made the computations on the sequence of parameters \( \alpha_i = r^{-i}, i = 0, 1, \ldots \) and \( r = 1.1 \). The parameter \( \alpha_i \) is found as the minimizer of the corresponding functions \( \tilde{g}_k(\alpha) \) (see (5.2)) in the interval \([\alpha, 1]\), where \( \alpha \) is the largest \( \alpha_i \) for which the value of \( \tilde{g}_k(\alpha_i) \) is \( C \) times larger than its value at its current minimum (see [8, 10, 24]). We used the value \( C = 1.2 \), in addition we found the first local minimum (corresponding to \( C = 1 \)) and the global minimum (corresponding to \( C = \infty \)). Table 4 shows the error ratios and success percent-
A family of rules in the Lavrentiev method

Figure 3

<table>
<thead>
<tr>
<th>$k$</th>
<th>$p$</th>
<th>$C = 1$</th>
<th>$C = 1.2$</th>
<th>$C = 1$</th>
<th>$C = 1.2$</th>
<th>$C = \infty$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>53.3</td>
<td>1.12</td>
<td>47%</td>
<td>99%</td>
<td>83%</td>
</tr>
<tr>
<td>0</td>
<td>2</td>
<td>48.0</td>
<td>1.06</td>
<td>83%</td>
<td>100%</td>
<td>81%</td>
</tr>
<tr>
<td>$1/2$</td>
<td>0</td>
<td>8.33</td>
<td>1.09</td>
<td>77%</td>
<td>100%</td>
<td>84%</td>
</tr>
<tr>
<td>$1/2$</td>
<td>2</td>
<td>1.03</td>
<td>1.03</td>
<td>100%</td>
<td>100%</td>
<td>81%</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>6.67</td>
<td>1.08</td>
<td>92%</td>
<td>100%</td>
<td>84%</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>1.02</td>
<td>1.02</td>
<td>100%</td>
<td>100%</td>
<td>81%</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>1.11</td>
<td>1.07</td>
<td>100%</td>
<td>100%</td>
<td>84%</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>1.02</td>
<td>1.02</td>
<td>100%</td>
<td>100%</td>
<td>81%</td>
</tr>
</tbody>
</table>

Table 4. Results in heuristic rules.

ages, i.e. the percentage of cases, where the error ratio was lower than 2. In case $k = 0$ the formulas (5.1) and (5.2) coincide. In formula (5.1) the parameter value $k > 0$ gave worse results than $k = 0$, with success percentages between 30% and 60% in all cases ($C = 1$, $C = 1.2$, $C = \infty$). The last column shows that searching the global minimum worked well in most cases but we don’t show the averages of error ratios, which were large due to failures in some runs.
Bibliography


Received September 1, 2012.

Author information
Uno Hämarik, Faculty of Mathematics and Computer Science, University of Tartu, J. Liivi 2, 50409 Tartu, Estonia.
E-mail: uno.hamarik@ut.ee

Reimo Palm, Faculty of Mathematics and Computer Science, University of Tartu, J. Liivi 2, 50409 Tartu, Estonia.
E-mail: reimo.palm@ut.ee

Toomas Raus, Faculty of Mathematics and Computer Science, University of Tartu, J. Liivi 2, 50409 Tartu, Estonia.
E-mail: toomas.raus@ut.ee