mathematics

Article

Q-Curve and Area Rules for Choosing Heuristic Parameter in Tikhonov Regularization

Toomas Raus and Uno Hämarik *

Institute of Mathematics and Statistics, University of Tartu, 51009 Tartu, Estonia; toomas.raus@ut.ee
* Correspondence: uno.hamarik@ut.ee

Received: 8 June 2020; Accepted: 12 July 2020; Published: 16 July 2020

Abstract: We consider choice of the regularization parameter in Tikhonov method if the noise level of the data is unknown. One of the best rules for the heuristic parameter choice is the quasi-optimality criterion where the parameter is chosen as the global minimizer of the quasi-optimality function. In some problems this rule fails. We prove that one of the local minimizers of the quasi-optimality function is always a good regularization parameter. For the choice of the proper local minimizer we propose to construct the Q-curve which is the analogue of the L-curve, but on the x-axis we use modified discrepancy instead of discrepancy and on the y-axis the quasi-optimality function instead of the norm of the approximate solution. In the area rule we choose for the regularization parameter such local minimizer of the quasi-optimality function for which the area of the polygon, connecting on Q-curve this minimum point with certain maximum points, is maximal. We also provide a posteriori error estimates of the approximate solution, which allows to check the reliability of the parameter chosen heuristically. Numerical experiments on an extensive set of test problems confirm that the proposed rules give much better results than previous heuristic rules. Results of proposed rules are comparable with results of the discrepancy principle and the monotone error rule, if the last two rules use the exact noise level.

Keywords: ill-posed problem; Tikhonov regularization; unknown noise level; regularization parameter choice; heuristic rule; quasi-optimality function

MSC: 45A52; 65J20

---

1. Introduction

Let $A \in \mathcal{L}(H, F)$ be a linear bounded operator between real Hilbert spaces H, F. We are interested in finding the minimum norm solution u_* of the equation

$$Au = f_*, \qquad f_* \in \mathcal{R}(A), \tag{1}$$

where noisy data $f \in F$ are given instead of the exact data f_*. The range $\mathcal{R}(A)$ may be non-closed and the kernel $\mathcal{N}(A)$ may be non-trivial, so in general this problem is ill-posed. We consider the solution of the problem $Au = f$ by the Tikhonov method (see [1,2]) where the regularized solutions in cases of exact and inexact data have the corresponding forms

$$u_\alpha^+ = (\alpha I + A^*A)^{-1} A^* f_*, \qquad u_\alpha = (\alpha I + A^*A)^{-1} A^* f$$

and $\alpha > 0$ is the regularization parameter. Using the well-known estimate $\|u_\alpha - u_\alpha^+\| \leq \frac{1}{2}\alpha^{-1/2}\|f - f_*\|$ (see [1,2]) and notations

$$e(\alpha) := \|u_\alpha - u_*\|, \qquad e_1(\alpha) := \|u_\alpha^+ - u_*\| + \|u_\alpha - u_\alpha^+\|, \tag{2}$$

$$e_2(\alpha, \|f - f_*\|) := \left\|u_\alpha^+ - u_*\right\| + \frac{1}{2\sqrt{\alpha}} \|f - f_*\|,$$

we have the error estimates

$$e(\alpha) \leq e_1(\alpha) \leq e_2(\alpha, \|f - f_*\|). \tag{3}$$

We consider choice of the regularization parameter if the noise level for $\|f - f_*\|$ is unknown. The parameter choice rules which do not use the noise level information are called heuristic rules. Well known heuristic rules are the quasi-optimality criterion [3–9], L-curve rule [10,11], GCV (generalized cross-validation)-rule [12], Hanke–Raus rule [13], Reginska's rule [14] and its modifications [15]; for other rules see [16–18]. Heuristic rules are numerically compared in [4,17–19]. The heuristic rules give good results in many problems, but it is not possible to construct heuristic rules guaranteeing convergence $\|u_\alpha - u_*\| \to 0$ as the noise level goes to zero (see [20]). All heuristic rules may fail in some problems and without additional information about the solution, it is difficult to decide, is the obtained parameter reliable or not.

In the quasi-optimality criterion the parameter α is chosen as the global minimizer of the function $\psi_Q(\alpha) = \alpha \left\| \frac{du_\alpha}{d\alpha} \right\|$ on certain interval. We propose to choose the parameter from the set $L_{\min}$ of local minimizers of this function from certain set Ω of parameters.

We will call the parameter α_R in arbitrary rule R as pseudooptimal, if

$$\|u_{\alpha_R} - u_*\| \leq c \min_{\alpha > 0} e_1(\alpha)$$

with relatively small constant c and we show that at least one parameter from set $L_{\min}$ has this property. For the choice of proper parameter from the set $L_{\min}$ some algorithms were proposed in [21], in the current work we propose other algorithms. We propose to construct the Q-curve which is the analogue of the L-curve [10], but on the x-axis we use modified discrepancy instead of discrepancy and on the y-axis function $\psi_Q(\alpha)$ instead of $\|u_\alpha\|$. For finding proper local minimizer of the function $\psi_Q(\alpha)$, we propose the area rules on the Q-curve. The idea of proposed rules is that we form, for every minimizer of the function $\psi_Q(\alpha)$, a certain function which approximates the error of the approximate solution and has one minimizer; we choose for the regularization parameter such local minimizer of $\psi_Q(\alpha)$, for which the area of the polygon, connecting this minimum point with certain maximum points, is maximal.

The plan of this paper is as follows. In Section 2 we consider known rules for choice of the regularization parameter, both in case of known and unknown noise level. In Section 3, we prove that the set $L_{\min}$ contains at least one pseudooptimal parameter. In Section 4, information about used test problems (mainly from [11,22], but also from [11,23–26]) and numerical experiments is given. In Section 5, we consider the Q-curve and area rule, in Section 6 further developments of the area rule. These algorithms are also illustrated by the results of numerical experiments.

2. Rules for the Choice of the Regularization Parameter

2.1. Parameter Choice in the Case of Known Noise Level

In the case of known noise level δ, $\|f - f_*\| \leq \delta$ we use one of the so-called δ-rules, where certain functional $d(\alpha)$ and constant $b \geq b_0$ (b_0 depends on $d(\alpha)$) are chosen and such regularization parameter $\alpha(\delta)$ is chosen which satisfies $d(\alpha) = b\delta$.

(1) Discrepancy principle (DP) [2,27]:

$$d_D(\alpha) := \|Au_\alpha - f\| = b\delta, \quad b \geq 1.$$

(2) Modified discrepancy principle (Raus–Gfrerer rule) [28,29]:

$$d_{\mathrm{MD}}(\alpha) := \|B_\alpha (Au_\alpha - f)\| = b\delta, \quad B_\alpha := \alpha^{1/2} (\alpha I + AA^*)^{-1/2}, \quad b \geq 1.$$

(3) Monotone error rule (ME-rule) [30,31]:

$$d_{\mathrm{ME}}(\alpha) := \frac{\|B_\alpha (Au_\alpha - f)\|^2}{\|B_\alpha^2 (Au_\alpha - f)\|} = \delta.$$

The name of this rule is justified by the fact that the chosen parameter α_{ME} satisfies

$$\frac{\mathrm{d}}{\mathrm{d}\alpha}\|u_\alpha - u_*\| > 0 \quad \forall \alpha > \alpha_{\mathrm{ME}}.$$

Therefore $\alpha_{\mathrm{ME}} \geq \alpha_{\mathrm{opt}} := \mathrm{argmin}\|u_\alpha - u_*\|$.

(4) Monotone error rule with post-estimation (MEe-rule) [4,18,32–35]. The inequality $\alpha_{\mathrm{ME}} \geq \alpha_{\mathrm{opt}}$ suggests to use somewhat smaller parameter than α_{ME}. Extensive numerical experiments suggest to compute α_{ME} and to use the post-estimated parameter $\alpha_{\mathrm{MEe}} := 0.4\alpha_{\mathrm{ME}}$. Then typically $\|u_{\alpha_{\mathrm{MEe}}} - u_*\| / \|u_{\alpha_{\mathrm{ME}}} - u_*\| \in (0.7, 0.9)$. If the exact noise level is known, this MEe-rule gives typically the best results from all δ-rules.

(5) Rule R1 [36]. Let $b > \frac{2}{3\sqrt{3}}$. Choose $\alpha(\delta)$ as the smallest solution of the equation

$$d_{\mathrm{R1}}(\alpha(\delta)) := \alpha^{-1/2} \left\| A^* B_\alpha^2 (Au_\alpha - f) \right\| = b\delta.$$

Note that this equation can be rewritten using the 2-iterated Tikhonov approximation $u_{2,\alpha}$:

$$B_\alpha^2 (Au_\alpha - f) = Au_{2,\alpha} - f, \qquad u_{2,\alpha} = (\alpha I + A^*A)^{-1} (\alpha u_\alpha + A^* f). \tag{4}$$

The last four rules are weakly quasioptimal rules (see [37]) for the Tikhonov method: if $\|f - f_*\| \leq \delta$, then $\left\| u_{\alpha(\delta)} - u_* \right\| \leq C(b) \inf_{\alpha > 0} e_2(\alpha, \delta)$ (see (3)). The rules for the parameter choice in case of approximately given noise level are proposed and analyzed in [18,32–34].

2.2. Parameter Choice in the Case of Unknown Noise Level

A classical heuristic rule is the quasi-optimality criterion. In the Tikhonov method, it chooses $\alpha = \alpha_{\mathrm{Q}}$ or $\alpha = \alpha_{\mathrm{QD}}$ as the global minimizer of corresponding functions

$$\psi_{\mathrm{Q}}(\alpha) = \alpha \left\| \frac{du_\alpha}{d\alpha} \right\| = \alpha^{-1} \left\| A^* B_\alpha^2 (Au_\alpha - f) \right\| = \alpha \|A^* (\alpha I + AA^*)^{-2} f\| = \|u_{2,\alpha} - u_\alpha\|, \tag{5}$$

$$\psi_{\mathrm{QD}}(\alpha) = (1 - q)^{-1} \left\| u_\alpha - u_{q\alpha} \right\|, \quad 0 < q < 1. \tag{6}$$

The Hanke–Raus rule finds parameter $\alpha = \alpha_{HR}$ as the global minimizer of the function

$$\psi_{\mathrm{HR}}(\alpha) = \alpha^{-1/2} \|B_\alpha (Au_\alpha - f)\|.$$

In practice, often the L-curve is used. The L-curve is the log-log-plot of $\|u_\alpha\|$ versus $\|Au_\alpha - f\|$. The points $(\|Au_\alpha - f\|, \|u_\alpha\|)$ have often a shape similar to the letter L and the parameter α_L which corresponds to the corner point is often a good parameter. In the literature several concrete rules for choice of the corner point are proposed. In [14], a parameter is chosen as the global minimizer of the function

$$\psi_{\mathrm{RE}}(\alpha) = \|Au_\alpha - f\| \, \|u_\alpha\|^\tau, \quad \tau \geq 1$$

(below we use this rule with $\tau = 1$). Another rule for choice of the corner point is the maximum curvature method ([38,39]), where such parameter α is chosen for which the curvature of the L-curve as the function

$$\psi_{\mathrm{MC}}(\alpha) = 2\frac{\hat{\rho}'\hat{\xi}'' - \hat{\rho}''\hat{\xi}'}{((\hat{\rho}')^2 + (\hat{\xi}')^2)^{3/2}}$$

is maximal. Here $\hat{\rho}', \hat{\xi}', \hat{\rho}'', \hat{\xi}''$ are first and second order derivatives of functions $\log d_{\mathrm{D}}(\alpha)$ and $\log \|u_\alpha\|$.

We propose also a new heuristic rule, where the global minimizer of the function

$$\psi_{\mathrm{WQ}}(\alpha) = d_{\mathrm{MD}}(\alpha)\psi_{\mathrm{Q}}(\alpha) \tag{7}$$

is chosen for the parameter. We call this rule as the weighted quasioptimality criterion.

In the following we will find the regularization parameter from the set of parameters

$$\Omega = \left\{\alpha_j : \alpha_j = q\alpha_{j-1}, \quad j = 1, 2, ..., N, \quad 0 < q < 1\right\}, \tag{8}$$

where α_0, q, α_N are given. If in the discretized problem the minimal eigenvalue $\lambda_{\min}$ of the matrix $A^T A$ is larger than α_N, the heuristic rules above often choose parameter α_N, which is generally not a good parameter. The works [6–8] propose to search the global minimum of the function $\psi_{\mathrm{Q}}(\alpha)$ in the interval $[\max(\alpha_N, \lambda_{\min}), \alpha_0]$.

Definition 1. *We say that the discretized problem $Au = f$ does not need regularization if*

$$e_1(\lambda_{\min}) \leq 2 \min_{\alpha \in \Omega,\ \alpha \geq \lambda_{\min}} e_1(\alpha).$$

If the discretized problem does not need regularization then $\alpha' = 0$ or $\alpha' = \alpha_N$ is the proper parameter while for $\alpha' \leq \lambda_{min}$ we have

$$\|u_{\alpha'} - u_*\| \leq e_1(\alpha') = \left\|u^+_{\alpha'} - u_*\right\| + \left\|(\alpha' I + A^*A)^{-1}A^*(f - f_*)\right\| \leq$$
$$\left\|u^+_{\lambda_{min}} - u_*\right\| + 2\left\|(\lambda_{min} I + A^*A)^{-1}A^*(f - f_*)\right\| \leq 2e_1(\lambda_{min}) \leq 4 \min_{\alpha \in \Omega, \alpha \geq \lambda_{min}} e_1(\alpha).$$

Searching the parameter from the interval $[\max(\alpha_N, \lambda_{\min}), \alpha_0]$ means the a priori assumption that the discretized problem needs regularization. Note that if $\lambda_{\min} > \alpha_N$, then in general it is not possible to decide (without additional information about solution or about noise of the data) whether the discretized problem needs regularization or not.

3. Local Minimum Points of the Function $\psi_{\mathrm{Q}}(\alpha)$

In the following we investigate the function $\psi_{\mathrm{Q}}(\alpha)$ in (5) and show that at least one local minimizer of this function is the pseudooptimal parameter. We need some preliminary results.

Lemma 1. *The functions $\psi_{\mathrm{Q}}(\alpha)$, $\psi_{\mathrm{QD}}(\alpha)$ satisfy for each $\alpha > 0$ the estimates*

$$\psi_{\mathrm{Q}}(\alpha) \leq e_1(\alpha), \tag{9}$$

$$\psi_{\mathrm{QD}}(\alpha) \leq q^{-1} e_1(\alpha), \tag{10}$$

$$\psi_{\mathrm{Q}}(\alpha) \leq \psi_{\mathrm{QD}}(\alpha) \leq q^{-1}\psi_{\mathrm{Q}}(q\alpha).$$

Proof. Using relations $f = Au_* + f - f_*$,

$$u_\alpha - u_{q\alpha} = (q-1)\,\alpha\,(\alpha I + A^*A)^{-1}\,(q\alpha I + A^*A)^{-1}\,A^*f,$$

$$\left\|A^*A\,(\alpha I + A^*A)^{-1}\right\| \leq 1, \quad \alpha\left\|(\alpha I + A^*A)^{-1}\right\| \leq 1,$$

we have

$$\begin{aligned}
\psi_{\mathrm{Q}}(\alpha) &= \alpha\|A^*(\alpha I+AA^*)^{-2}f\| = \alpha\|(\alpha I+A^*A)^{-2}A^*f\| \\
&\le \alpha\left\|A^*A(\alpha I+A^*A)^{-2}u_*\right\| + \alpha\left\|(\alpha I+A^*A)^{-2}A^*(f-f_*)\right\| \\
&\le \alpha\|(\alpha I+A^*A)^{-1}u_*\| + \|(\alpha I+A^*A)^{-1}A^*(f-f_*)\| = e_1(\alpha), \\
\psi_{\mathrm{QD}}(\alpha) &\le \alpha\left\|A^*A(q\alpha I+A^*A)^{-1})(\alpha I+A^*A)^{-1})u_*\right\| \\
&\quad +\alpha\left\|(q\alpha I+A^*A)^{-1})(\alpha I+A^*A)^{-1})A^*(f-f_*)\right\| \le q^{-1}e_1(\alpha), \\
\psi_{\mathrm{Q}}(\alpha) &= \alpha\left\|(\alpha I+A^*A)^{-2}A^*f\right\| \le \alpha\left\|(\alpha I+A^*A)^{-1}(q\alpha I+A^*A)^{-1}A^*f\right\| \\
&= \psi_{\mathrm{QD}}(\alpha) \le \alpha\left\|(q\alpha I+A^*A)^{-2}A^*f\right\| = q^{-1}\psi_{\mathrm{Q}}(q\alpha).
\end{aligned}$$

□

Remark 1. *Note that* $\lim_{\alpha\to\infty}\psi_{\mathrm{Q}}(\alpha)=0$, *but* $\lim_{\alpha\to\infty}e_1(\alpha)=\|u_*\|$. *Therefore in the case of too large* α_0 *this* α_0 *may be a global (or local) minimizer of the function* $\psi_{\mathrm{Q}}(\alpha)$. *The scaling argument suggests: by multiplying the equation* $A^*Au=A^*f$ *by some constant, it is also necessary to multiply* α *by this constant in the Tikhonov method. Therefore the parameter* α_0 *should be proportional to* $\|A^*A\|$. *We recommend to take* $\alpha_0=c\,\|A^*A\|$, $c\le 1$ *or to minimize the function* $\overline{\psi}_{\mathrm{Q}}(\alpha):=(1+\alpha/\,\|A^*A\|)\psi_{\mathrm{Q}}(\alpha)$ *instead of* $\psi_{\mathrm{Q}}(\alpha)$. *Due to the equality* $\lim_{\alpha\to 0}(1+\alpha/\,\|A^*A\|)=1$, *the function* $\overline{\psi}_{\mathrm{Q}}(\alpha)$ *approximately satisfies (9) for small* α.

In the following we define the local minimum points of the function $\psi_{\mathrm{Q}}(\alpha)$ on the set Ω (see (8)).

Definition 2. *We say that the parameter* α_k, $0\le k\le N-1$ *is the local minimum point of the sequence* $\psi_{\mathrm{Q}}(\alpha_k)$, *if* $\psi_{\mathrm{Q}}(\alpha_k)<\psi_{\mathrm{Q}}(\alpha_{k+1})$ *and in case* $k>0$ *there exists index* $j\ge 1$ *such that* $\psi_{\mathrm{Q}}(\alpha_k)=\psi_{\mathrm{Q}}(\alpha_{k-1})=...=\psi_{\mathrm{Q}}(\alpha_{k-j+1})<\psi_{\mathrm{Q}}(\alpha_{k-j})$. *The parameter* α_N *is the local minimum point if there exists index* $j\ge 1$ *so that*

$$\psi_{\mathrm{Q}}(\alpha_N)=\psi_{\mathrm{Q}}(\alpha_{N-1})=...=\psi_{\mathrm{Q}}(\alpha_{N-j+1})<\psi_{\mathrm{Q}}(\alpha_{N-j}).$$

Denote the local minimum points by m_k, $k=1,\dots,K$ *(*K *is the number of minimum points) and corresponding set by* $L_{\min}=\{m_k : m_1>m_2>...>m_K\}$.

Definition 3. *The parameter* α_k, $0<k<N$ *is the local maximum point of the sequence* $\psi_{\mathrm{Q}}(\alpha_k)$ *if* $\psi_{\mathrm{Q}}(\alpha_k)>\psi_{\mathrm{Q}}(\alpha_{k+1})$ *and there exists index* $j\ge 1$ *so that*

$$\psi_{\mathrm{Q}}(\alpha_k)=\psi_{\mathrm{Q}}(\alpha_{k-1})=...=\psi_{\mathrm{Q}}(\alpha_{k-j+1})>\psi_{\mathrm{Q}}(\alpha_{k-j}).$$

We denote by M_k *the local maximum point between the local minimum points* m_{k+1} *and* m_k, $1\le k\le K-1$. *Denote* $M_0=\alpha_0$, $M_K=\alpha_N$. *Then by the construction*

$$M_K\le m_K<M_{K-1}<\ldots<m_2<M_1<m_1\le M_0.$$

Theorem 1. *The following estimates hold for the local minimizers of the function* $\psi_{\mathrm{Q}}(\alpha)$.

1. *For arbitrary* α_0,α_N *we have*

$$\min_{\alpha\in L_{\min}}\|u_\alpha-u_*\|\le q^{-1}C\min_{\alpha_N\le\alpha\le\alpha_0}e_1(\alpha), \tag{11}$$

$$C:=1+\max_{1\le k\le K}\ \max_{\alpha_j\in\Omega,\,M_k\le\alpha_j\le M_{k-1}}T\left(m_k,\alpha_j\right)\le 1+c_q\ln\left(\frac{\alpha_0}{\alpha_N}\right),\ T(\alpha,\beta):=\frac{\left\|u_\alpha-u_\beta\right\|}{\psi_{\mathrm{Q}}(\beta)}.$$

2. *If* $\alpha_0 = \|A^*A\|$, $\alpha_N = \alpha_0 \left(\frac{\|f-f_*\|}{2\|u_*\|}\right)^2$, *then*

$$\min_{\alpha \in L_{\min}} \|u_\alpha - u_*\| \leq q^{-1}(1 + 2\max\{1, c_q \mid \ln \frac{\|f - f_*\|}{2\,\|A\|\,\|u_*\|} \mid\}) \min_{\alpha>0} e_2(\alpha, \|f - f_*\|), \tag{12}$$

where $c_q := \left(q^{-1} - 1\right) / \ln q^{-1} \to 1$ *if* $q \to 1$.

Moreover, if $u_* = |A|^p\, v$, $\|v\| \leq \rho$, $p > 0$, *where* $|A| := (A^*A)^{1/2}$, *then*

$$\min_{\alpha \in L_{\min}} \|u_\alpha - u_*\| \leq c_{p,q}\rho^{\frac{1}{p+1}} \left|\ln \|f - f_*\|\right| \|f - f_*\|^{\frac{p}{p+1}}, 0 < p \leq 2. \tag{13}$$

Proof. For arbitrary parameters $\alpha \geq 0,\ \beta \geq 0$ the inequalities

$$\|u_\alpha - u_*\| \leq \left\|u_\alpha - u_\beta\right\| + \left\|u_\beta - u_*\right\| \leq T(\alpha, \beta)\psi_Q(\beta) + e_1(\beta)$$

and (9) lead to the estimate

$$\|u_\alpha - u_*\| \leq (1 + T(\alpha, \beta))\, e_1(\beta). \tag{14}$$

It is easy to see that

$$\min_{\alpha_j \in \Omega} e_1(\alpha_j) \leq q^{-1} \min_{\alpha_N \leq \alpha \leq \alpha_0} e_1(\alpha), \tag{15}$$

while in case $q\alpha \leq \alpha' \leq \alpha$ we have $e_1\,(\alpha') \leq q^{-1} e_1\,(\alpha)$.

Let $\alpha_{j*} = \alpha_0 q^{j*}$ be the global minimizer of the function $e_1(\alpha)$ on the set of the parameters Ω. Then, $\alpha_{j*} \in [M_k, M_{k-1}]$ for some $k,\ 1 \leq k \leq K$ and this k defines index m with $m_k = \alpha_m$. From (14) we get the estimate

$$\left\|u_{m_k} - u_*\right\| \leq \left(1 + T(m_k, \alpha_{j*})\right) e_1(\alpha_{j*}) \leq \left(1 + \min_{M_k \leq \alpha_j \leq M_{k-1}} T(m_k, \alpha_j)\right) \min_{\alpha_j \in \Omega} e_1(\alpha_j),$$

which together with (15) gives also the estimate (11).

Now we show that $C \leq 1 + c_q \ln\left(\frac{\alpha_0}{\alpha_N}\right)$. If $m_k \leq \alpha_j \leq M_{k-1}$, then using Lemma 1 and equality (6) we get

$$\left\|u_{\alpha_m} - u_{\alpha_j}\right\| \leq \sum_{j \leq i \leq m-1} \|u_i - u_{i+1}\| \leq q^{-1}(1-q) \sum_{j \leq i \leq m-1} \psi_Q(\alpha_{i+1}).$$

Due to the inequalities $\psi_Q(\alpha_{i+1}) \leq \psi_Q(\alpha_j)$ for all $i, j \leq i \leq m-1$ we have

$$\begin{aligned} T(m_k, \alpha_j) &= \frac{\left\|u_{\alpha_m} - u_{\alpha_j}\right\|}{\psi_Q(\alpha_j)} \leq q^{-1}(1-q) \sum_{j \leq i \leq m-1} \frac{\psi_Q(\alpha_{i+1})}{\psi_Q(\alpha_j)} \\ &\leq (q^{-1} - 1)(m - j) \leq (q^{-1} - 1)N = \frac{(q^{-1} - 1)}{\ln q^{-1}} \ln \frac{\alpha_0}{\alpha_N} = c_q \ln \frac{\alpha_0}{\alpha_N}. \end{aligned}$$

If $M_k \leq \alpha_j \leq m_k$, then analogous estimation of $T(m_k, \alpha_j)$ gives the same result.

Now we prove the estimate (12). For the global minimum point α_* of the function $e_2(\alpha, \|f - f_*\|)$ the inequality $\alpha_* \geq \alpha_N$ holds, while for $\alpha < \alpha_N$ we have

$$e_2(\alpha^*) \leq \|u_*\| = \|f - f_*\| / (2\sqrt{\alpha_N}) \leq \|f - f_*\| / (2\sqrt{\alpha}) < e_2(\alpha).$$

In the case $\alpha_* \leq \alpha_0$ we get similarly as in the proof of the estimate (11) that

$$\min_{\alpha \in L_{\min}} \|u_\alpha - u_*\| \leq q^{-1}(1 + c_q \ln \frac{\alpha_0}{\alpha_N}) \min_{\alpha > 0} e_2(\alpha, \|f - f_*\|);$$

due to $\ln \frac{\alpha_0}{\alpha_N} = |\ln \|f - f_*\| / \|f_*\| \, |$, the estimate (12) holds. Consider the case $\alpha_* > \alpha_0$. Then,

$$e_2(\alpha_*, \|f - f_*\|) \geq \left\|u^+_{\alpha_0} - u_*\right\| \geq \frac{\alpha_0}{\alpha_0 + \|A^* A\|} \|u_*\| = \frac{\|u_*\|}{2}$$

and for each local minimum point $m_k, \alpha_N \leq m_k \leq \alpha_0$ the inequalities

$$\begin{aligned}\left\|u_{m_k} - u_*\right\| \leq e_2(m_k, \|f - f_*\|) \leq \left\|u^+_{\alpha_0} - u_*\right\| + 0.5{\alpha_N}^{-1/2} \|f - f_*\| = \\ \left\|u^+_{\alpha_0} - u_*\right\| + \|u_*\| \leq 3 \left\|u^+_{\alpha_0} - u_*\right\| \leq 3e_2(\alpha_*, \|f - f_*\|)\end{aligned}$$

hold. Therefore the inequality (12) holds also in this case.

For source-like solution $u_* = |A|^p v$, $\|v\| \leq \rho$, $p > 0$ the error estimate

$$\min_{\alpha_N \leq \alpha \leq \alpha_0} e_1(\alpha) \leq c_p \rho^{1/(p+1)} \|f - f_*\|^{p/(p+1)}, 0 < p \leq 2$$

is well-known (see [1,2]) and the estimate (13) follows immediately from (12). □

Remark 2. *Theorem 3 holds also in the case if the equation $Au = f_*$ has only the quasisolution, i.e., in the case $f_* \notin \mathcal{R}(A)$, $Qf_* \in \mathcal{R}(A)$, where Q is the orthoprojector $F \to \overline{\mathcal{R}(A)}$.*

Remark 3. *The inequality (11) holds also in the case if the noise of the function f is not finite but $\min_{\alpha_N \leq \alpha \leq \alpha_0} e_1(\alpha)$ is finite (this holds if $\|A^*(f - f_*)\|$ is finite).*

Remark 4. *Use of the inequality (10) enables to prove the analogue of Theorem 3 for set $L_{\min}$ of local minimizers of the function $\psi_{\text{QD}}(\alpha)$: then the inequality (11) holds, where $T(\alpha, \beta) = q^{-1} \frac{\|u_\alpha - u_\beta\|}{\psi_{QD}(\beta)}$.*

In choice of the regularization parameter we may exclude from the observation some local minimizers. It is natural to assume that α_N is so small that

$$d_{\text{MD}}(\alpha_N) \leq (1 + \epsilon) \|f - f_*\| \tag{16}$$

with small $\epsilon > 0$. Then the following theorem holds.

Theorem 2. *Let (16) hold. Let m_{k_0} be some local minimizer in $L_{\min}$. Then*

$$\min_{\alpha \in L_{\min}, \alpha \geq m_{k_0}} \|u_\alpha - u_*\| \leq \max\{q^{-1} C_1 \min_{\alpha \geq 0} e_1(\alpha), C_2(b, \epsilon) \min_{\alpha \geq 0} e_2(\alpha, \|f - f_*\|)\},$$

where $b = d_{\text{MD}}(m_{k_0}) / d_{\text{MD}}(\alpha_N) \geq 1$, $\quad C_2(b, \epsilon) := b(1 + \epsilon) + 2$ and

$$C_1 := 1 + \max_{1 \leq k \leq k_0} \max_{\alpha_j \in \Omega,\, M_k \leq \alpha_j \leq M_{k-1}} T\left(m_k, \alpha_j\right) \leq 1 + c_q \ln\left(\frac{\alpha_0}{m_{k_0}}\right).$$

Proof. Let α_i^*, $i = 1, 2$ be global minimizers of the functions $e_1(\alpha)$ and $e_2(\alpha, \|f - f_*\|)$, respectively. We consider separately 3 cases. If $m_{k_0} \leq \alpha_1^*$ we get similarly to the proof of Theorem 3 the estimate

$$\min_{\alpha \in L_{\min}, \alpha \geq m_{k_0}} \|u_\alpha - u_*\| \leq q^{-1} C_1 \min_{\alpha_N \leq \alpha \leq \alpha_0} e_1(\alpha). \tag{17}$$

If $\alpha_1^* \le m_{k_0} < \alpha_2^*$, we estimate

$$\left\| u_{m_{k_0}} - u_* \right\| \le \left\| u^+_{\alpha_2^*} - u_* \right\| + \frac{\|f - f_*\|}{2\sqrt{\alpha_1^*}} \le \min_{\alpha \ge 0} e_2(\alpha, \|f - f_*\|) + \min_{\alpha \ge 0} e_1(\alpha). \quad (18)$$

If $\alpha_1^* \le m_{k_0}$ and $\alpha_2^* \le m_{k_0}$, we have

$$\left\| B_{m_{k_0}} \left(A u_{m_{k_0}} - f \right) \right\| \le b d_{\mathrm{MD}}(\alpha_N) \le b(1+\epsilon) \|f - f_*\|$$

and now we can prove analogically to the proof of the weak quasioptimality of the modified discrepancy principle ([37]) that under assumption $\alpha_2^* \le m_{k_0}$ the error estimate

$$\left\| u_{m_{k_0}} - u_* \right\| \le C_2(b, \epsilon) \min_{\alpha \ge 0} e_2(\alpha, \|f - f_*\|) \quad (19)$$

holds. Now the assertion 1 of Theorem 7 follows from the inequalities (17)–(19). □

4. On Test Problems and Numerical Experiments

We made numerical experiments for the local minimizers of the function $\psi_Q(\alpha)$ using three sets of test problems. The first set contains 10 well-known test problems from Regularization Toolbox [11] and the following 6 Fredholm integral equations of the first kind (discretized by the midpoint quadrature formula)

$$\int_a^b K(t,s) u(s) ds = f(t), \quad c \le t \le d.$$

- Groetsch1 [24]: $K(t,s) = \frac{t \exp(-t^2/(4s))}{2\sqrt{\pi} s^{3/2}},\ 0 \le s,t \le 100,\ u(s) = 40 + 5\cos((100-s)/5) + 2.5\cos(2(100-s)/2.5) + 1.25\cos(4(100-s)/2);$
- Groetsch2 [24]: $K(t,s) = \sum_{1 \le k \le 100} \frac{\sin(kt)\sin(ks)}{k},\ 0 \le s,t \le \pi,\ u(s) = s(\pi - s);$
- Indram [25] : $K(t,s) = e^{-st},\ 0 \le s,t \le 1,\ u(s) = s,\ f(t) = \frac{1-(t+1)e^{-t}}{t^2};$
- Ursell [11]: $K(t,s) = \frac{1}{1+s+t},\ 0 \le s,t \le 1,\ u(s) = s(1-s),\ f(t) = \frac{3+2t}{2} + (2+3t+t^2)\log\left(\frac{1+t}{2+t}\right);$
- Waswaz [26]: $K(t,s) = \cos(t-s),\ 0 \le s,t \le \pi,\ u(s) = \cos(s),\ f(t) = \frac{\pi}{2}\cos(t);$
- Baker [23]: $K(t,s) = e^{st},\ 0 \le s,t \le 1, u(s) = e^s,\ f(t) = \frac{e^{t+1}-1}{t+1}.$

The second set of test problems are well-known problems from [22]: Gauss, Hilbert, Lotkin, Moler, Pascal, Prolate. As in [22], we combined these six $n \times n$ matrices with 6 solution vectors $x_i = 1, x_i = i/n, x_i = ((i - [n/2])/[n/2])^2, x_i = \sin(2\pi(i-1)/n), x_i = i/n + 1/4\sin(2\pi(i-1)/n)$, $x_i = 0$ if $i \le [n/2]$ and $x_i = 1$ if $i > [n/2]$. For getting the third set of test problems we combined the matrices of the first set of test problems with 6 solutions of the second set of test problems.

Numerical experiments showed that performance of different rules depends essentially on eigenvalues of the matrix $A^T A$. We characterize these eigenvalues via three indicators: the value of minimal eigenvalue $\lambda_{\min}$, by the value N_1, showing the number of eigenvalues less than α_N and by the value Λ, characterizing the density of location of eigenvalues on the interval $[\max(\alpha_N, \lambda_{\min}), 1]$. More precisely, let the eigenvalues of the matrix $A^T A$ be $\lambda_1 \ge \lambda_2 \ge ... \ge \lambda_n = \lambda_{\min}$. Then the value of Λ is found by the formula $\Lambda = \max_{\lambda_k > \max(\alpha_N, \lambda_n)} \lambda_k / \lambda_{k+1}$. We characterize the smoothness of the solution by the value

$$p1 = \frac{\log \min_\alpha e_2(\alpha, \|f - f_*\|) - \log \|u_*\|}{\log \|f - f_*\| - \log \|f_*\|},$$

where $\|f - f_*\| = 10^{-6}$. Table 1 contains the results of characteristics of the matrix $A^T A$ in case $n = 100,\ \alpha_N = 10^{-18}$.

Table 1. Characteristics of matrix A^TA and the solution u_*.

Problem	$\lambda_{\min}$	N_1	Λ	p1	Problem	$\lambda_{\min}$	N_1	Λ	p1
Baart	5.2×10^{-35}	92	1665.7	0.197	Spikes	1.3×10^{-33}	89	1529.3	0.005
Deriv2	6.7×10^{-9}	0	16.0	0.286	Wing	2.9×10^{-37}	94	9219.1	0.057
Foxgood	9.0×10^{-33}	85	210.1	0.426	Baker	1.0×10^{-33}	94	9153.1	0.498
Gravity	1.6×10^{-33}	68	4.1	0.403	Ursell	6.9×10^{-34}	94	3090.2	0.143
Heat	5.5×10^{-33}	3	2.4×10^{20}	0.341	Indramm	2.7×10^{-33}	94	9154.6	0.395
Ilaplace	3.8×10^{-33}	79	16.1	0.211	Waswaz2	2.0×10^{-34}	98	1.7×10^{30}	0.654
Phillips	1.4×10^{-13}	0	9.4	0.471	Groetsch1	5.8×10^{-33}	78	11.2	0.176
Shaw	2.3×10^{-34}	85	289.7	0.244	Groetsch2	1.0×10^{-4}	0	4.0	0.652

In all tests the discretization parameters $n \in \{60, 80, 100, 120, 140, 160, 180\}$ were used. We present the results of numerical experiments in tables for $n = 100$. Since the performance of rules generally depends on the smoothness p of the exact solution in (1), we complemented the standard solutions u_* of (now discrete) test problems with smoothened solutions $|A|^p u_*$, $p = 2$ computing the right-hand side as $A(|A|^p u_*)$. Results for $p = 2$ are given in Table 2, in all other tables and figures $p = 0$. After discretization, all problems were scaled (normalized) in such a way that the norms of the operator and the right-hand side were 1. All norms here and in the text below are Euclidean norms. On the base of exact data f_* we formed the noisy data f, where $\|f - f_*\|$ has values $10^{-1}, 10^{-2}, ..., 10^{-6}$, noise $f - f_*$ has normal distribution and the components of the noise were uncorrelated. We generated 20 noise vectors and used these vectors in all problems. We search the regularization parameter from the set Ω, where $\alpha_0 = 1, q = 0.95$ and N is chosen so that $\alpha_N \geq 10^{-18} > \alpha_{N+1}$. To guarantee that calculation errors do not influence essentially the numerical results, calculations were performed on geometrical sequence of decreasing α-s and finished for largest α with $d_{\mathrm{MD}}(q\alpha) > d_{\mathrm{MD}}(\alpha)$, while theoretically the function $d_{\mathrm{MD}}(\alpha)$ is monotonically increasing. Actually this precautionary measure was needed only in problem Groetsch2, calculations on $\alpha > \alpha_N$ were finished only in this problem. Since in model equations the exact solution is known, it is possible to find the regularization parameter α_*, which gives the smallest error on the set Ω. For every rule R the error ratio

$$E = \frac{\|u_{\alpha_{\mathrm{R}}} - u_*\|}{\|u_{\alpha_*} - u_*\|} = \frac{\|u_{\alpha_{\mathrm{R}}} - u_*\|}{\min_{\alpha \in \Omega} \|u_\alpha - u_*\|}$$

describes the performance of the rule R on this particular problem. To compare the rules or to present their properties, the following tables show averages and maximums of these error ratios over various parameters of the data set (problems, noise levels δ). We say that the heuristic rule fails if the error ratio $E > 100$. In addition to the error ratio E we present in some cases also error ratios

$$E1 = \frac{\|u_{\alpha_{\mathrm{R}}} - u_*\|}{\min_{\alpha \in \Omega} e_1(\alpha)}, \quad E2 = \frac{\|u_{\alpha_{\mathrm{R}}} - u_*\|}{\min_{\alpha \in \Omega} e_2(\alpha)}.$$

The results of numerical experiments for local minimizers $\alpha \in L_{\min}$ of the function $\psi_{\mathrm{Q}}(\alpha)$ are given in Table 3. For comparison, the results of δ-rules with $\delta = \|f - f_*\|$ are presented in the columns 2–4. Columns 5 and 6 contain respectively the averages and maximums of error ratios E for the best local minimizer $\alpha \in L_{\min}$. The results show that for many problems the Tikhonov approximation with the best local minimizer $\alpha \in L_{\min}$ is even more accurate than with the δ-rules parameters $\alpha_{\mathrm{ME}}, \alpha_{\mathrm{MEe}}$ or α_{DP}. Tables 1 and 3 show also that for rules ME and MEe the average error ratio E may be relatively large for problems where Λ is large and most of eigenvalues are smaller than α_N, while in this case $\min_{\alpha \in \Omega} e(\alpha)$ may be essentially smaller than $\min_{\alpha \in \Omega} e_2(\alpha, \|f_* - f\|)$. In these problems the discrepancy principle gives better parameter than ME and MEe rules. The average error in ME rule was largest in the problem Waswaz2, but error ratio E2 there is still under 1. This is due to the fact that for the tasks where the size of a Λ (defined on previous page) is large, the minimal error may be significantly smaller than $\min e_2(\alpha)$.

Table 2. Results of the numerical experiments, $p = 2$.

Problem	ME Aver E	MEe Aver E	DP Aver E	Best of $L_{\min}$ Aver E	$\|L_{\min}\|$ Aver	Combined Area Rule Aver E	Combined Area Rule Max E
Baart	1.86	1.19	2.93	1.18	4.74	1.60	14.57
Deriv2	1.09	1.19	3.65	1.03	2.00	1.04	1.17
Foxgood	1.56	1.13	3.58	1.14	2.08	1.22	3.58
Gravity	1.33	1.05	2.65	1.09	1.72	1.14	3.18
Heat	1.13	1.12	2.55	1.05	2.10	1.05	1.14
Ilaplace	1.47	1.06	2.78	1.11	2.73	1.13	3.51
Phillips	1.26	1.06	3.35	1.04	2.10	1.04	1.20
Shaw	1.37	1.06	2.58	1.11	3.72	1.29	8.96
Spikes	1.85	1.12	2.10	1.19	4.78	1.31	5.75
Wing	1.67	1.14	2.47	1.22	4.53	1.75	6.63
Baker	2.11	1.29	2.96	1.21	4.38	1.77	11.33
Ursell	1.86	1.19	4.10	1.16	4.82	1.67	18.08
Indramm	1.69	1.14	2.87	1.28	4.53	1.91	6.42
Waswaz2	127.2	49.8	1.20	2.44	1.00	2.43	9.01
Groetsch1	1.40	1.06	2.36	1.11	2.14	1.14	4.56
Groetsch2	1.02	1.23	1.71	1.14	1.67	1.55	3.97
Set 1	9.37	4.18	2.74	1.22	3.06	1.44	18.08
Set 2	2.10	1.26	2.91	1.19	2.83	1.37	29.03
Set 3	6.86	3.21	2.68	1.18	3.12	1.42	52.98

Table 3. Results for the set $L_{\min}$.

Problem	ME Aver E	MEe Aver E	DP Aver E	Best of $L_{\min}$ Aver E	Best of $L_{\min}$ Max E	$\|L_{\min}\|$ Aver	$\|L_{\min}\|$ Max	Apost. C Aver	Apost. C Max
Baart	1.43	1.32	1.37	1.23	2.51	6.91	8	3.19	3.72
Deriv2	1.29	1.07	1.21	1.08	1.34	1.71	2	3.54	4.49
Foxgood	1.98	1.42	1.34	1.47	6.19	3.63	6	3.72	4.16
Gravity	1.40	1.13	1.16	1.13	1.83	1.64	3	3.71	4.15
Heat	1.19	1.03	1.05	1.12	2.36	3.19	5	3.92	4.50
Ilaplace	1.33	1.21	1.26	1.20	2.56	2.64	5	4.84	6.60
Phillips	1.27	1.02	1.02	1.06	1.72	2.14	3	3.99	4.66
Shaw	1.37	1.24	1.28	1.19	2.15	4.68	7	3.48	4.43
Spikes	1.01	1.00	1.01	1.00	1.02	8.83	10	3.27	3.70
Wing	1.16	1.13	1.15	1.09	1.38	5.20	6	3.07	3.72
Baker	3.91	2.38	2.09	2.31	16.17	5.38	6	3.14	3.72
Ursell	2.14	1.97	2.03	1.69	4.44	5.53	6	3.07	3.43
Indramm	5.20	3.26	3.37	3.38	25.67	5.64	6	3.08	3.71
Waswaz2	127.2	49.9	1.20	2.44	9.03	1.00	1	2.00	2.00
Groetsch1	1.12	1.07	1.08	1.06	1.51	3.99	7	4.23	5.20
Groetsch2	1.02	1.22	1.67	1.13	1.69	1.67	2	5.62	13.72
Set 1	9.62	4.46	1.46	1.48	25.67	3.99	10	3.67	13.72
Set 2	1.57	1.32	1.36	1.20	5.33	4.40	10	3.50	5.43
Set 3	7.19	3.45	1.47	1.48	61.02	3.64	10	3.73	9.12

Columns 7 and 8 contain the averages and maximums of cardinalities $|L_{\min}|$ of sets $L_{\min}$ (number of elements of these sets). Note that the number of local minimizers depends on the parameter q (for smaller q the number of local minimizers is smaller) and on the length of minimization interval determined by the parameters α_N, α_0. The number of local minimizers is smaller also for larger noise level. Columns 9 and 10 contain the averages and maximums of values of constant C in the a posteriori error estimate (11). The value of C and error estimate (11) allow to assert, that in our test problems the choice of α as the best local minimizer in $L_{\min}$ guarantees that the error of the Tikhonov approximation has the same order as $\min_{\alpha_N \le \alpha \le \alpha_0} e_1(\alpha)$. Note that over all test problems the maximum of error ratio $E1$ for the best local minimizer in $L_{\min}$ and for the discrepancy principle were 1.93 and 9.90, respectively.

This confirms the result of Theorem 3 that at least one minimizer of the function $\psi_{\mathrm{Q}}(\alpha)$ is a good regularization parameter.

5. Q-Curve and Triangle Area Rule for Choosing Heuristic Regularization Parameter

We showed in the previous section that at least one local minimizer of the function $\psi_{\mathrm{Q}}(\alpha)$ is pseudooptimal parameter and we may omit small local minimizers α, for which $d_{\mathrm{MD}}(\alpha)$ is only slightly larger than $d_{\mathrm{MD}}(\alpha_N)$. We propose to construct for the parameter choice the Q-curve. The Q-curve figure uses the log-log scale with functions $d_{\mathrm{MD}}(\alpha)$ and $\psi_{\mathrm{Q}}(\alpha)$ on the x-axis and y-axis, respectively. The Q-curve can be considered as the analogue of the L-curve, where functions $Au_\alpha - f$ and $u_\alpha = -\alpha^{-1}A^*(Au_\alpha - f)$ are replaced by functions $B_\alpha(Au_\alpha - f)$ and $-\alpha^{-1}A^*B_\alpha^2(Au_\alpha - f)$ (see (4)), respectively. We denote $\tilde{d}_{\mathrm{MD}}(\alpha) := \log_{10} d_{\mathrm{MD}}(\alpha)$, $\tilde{\psi}_{\mathrm{Q}}(\alpha) := \log_{10} \psi_{\mathrm{Q}}(\alpha)$. For many problems the curve $(\tilde{d}_{\mathrm{MD}}(\alpha), \tilde{\psi}_{\mathrm{Q}}(\alpha))$ (or a part of this) has the form of letter L or V and we choose the minimizer at the "corner" point of L or V. We use the common logarithm instead of the natural logarithm, while then the Q-curve allows easier estimation of the supposed value of the noise level. In Figures 1–8 $n = 100$ is used, in Figures 9 and 10 $n = 60$. In Figures 1–4 the L-curves and Q-curves are compared for two problems, the global minimizer α_{opt} of the function $e_1(\alpha)$ is also presented. Note that in problem Baart $\lambda_{\min} < \alpha_N$ and in problem Deriv2 $\lambda_{\min} > \alpha_N$.

In most cases one can see on the Q-curve only one clear corner area with one local minimizer and then we take this local minimizer as corner point. If the corner area contains several local minimizers, we recommend to choose such local minimizer, for which the sum of coordinates of the corresponding point on the Q-curve is minimal. If Q-curve has several corner areas we recommend to use the very right of them. Actually, it is useful to present in the parameter choice besides the figures for every local minimizer m_k of the function $\psi_{\mathrm{Q}}(\alpha)$ also the coordinates of the point $(\tilde{d}_{\mathrm{MD}}(m_k), \tilde{\psi}_{\mathrm{Q}}(m_k))$ and the sums of the coordinates.

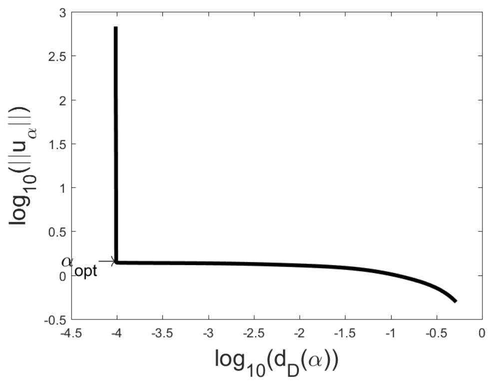

Figure 1. L-curve for Baart.

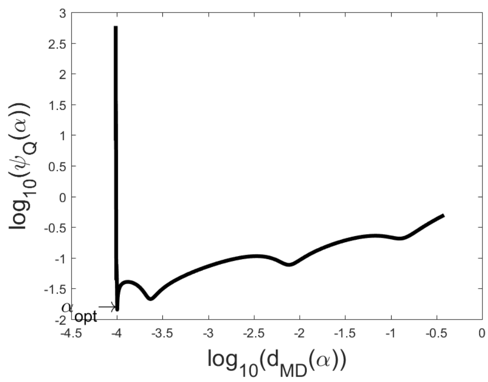

Figure 2. Q-curve for Baart.

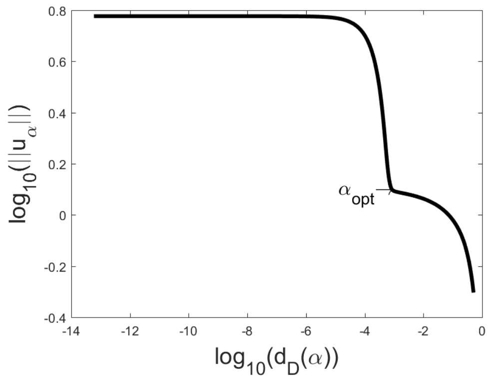

Figure 3. L-curve for Deriv2.

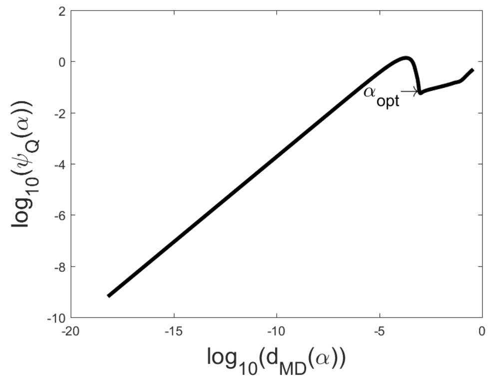

Figure 4. Q-curve for Deriv2.

For finding the proper local minimizer of the function $\psi_Q(\alpha)$ we present now a rule which works well for all test problems from set 1. The idea of the rule is to search the proper local minimizer m_k constructing certain triangles on the Q-curve and finding which of them has the maximal area. For every parameter α corresponds a point $P(\alpha)$ on the Q-curve with the corresponding coordinates $\left(\tilde{d}_{\text{MD}}(\alpha), \tilde{\psi}_Q(\alpha)\right)$. For every local minimizer m_k of the function $\psi_Q(\alpha)$ corresponds a triangle $T(k, r(k), l(k))$ with the vertices $P(m_k), P(M_{r(k)})$ and $P(M_{l(k)})$ on the Q-curve, where indices $r(k)$ and $l(k)$ correspond to the largest local maximums of the function $\psi_Q(\alpha)$ on two sides of the local minimum m_k:

$$\psi_Q(M_{r(k)}) = \max_{j<k} \psi_Q(M_j), \qquad \psi_Q(M_{l(k)}) = \max_{j \geq k} \psi_Q(M_j).$$

Triangle area rule (TA-rule). We choose for the regularization parameter such local minimizer m_k of the function $\psi_Q(\alpha)$ for which the area of the triangle $T(k, r(k), l(k))$ is the largest.

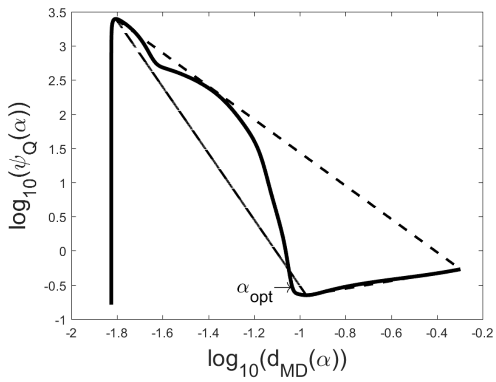

Figure 5. Q-curve in Heat.

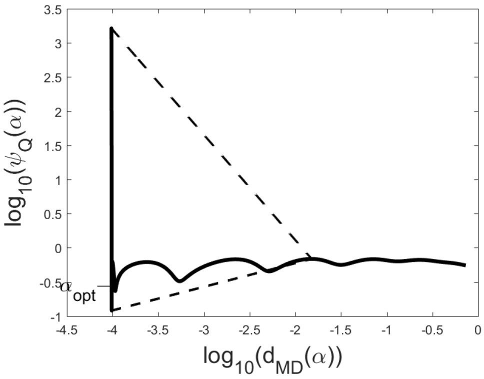

Figure 6. Q-curve in Spikes.

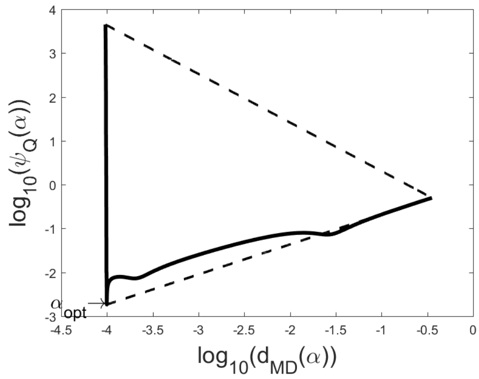

Figure 7. Q-curve in Foxgood.

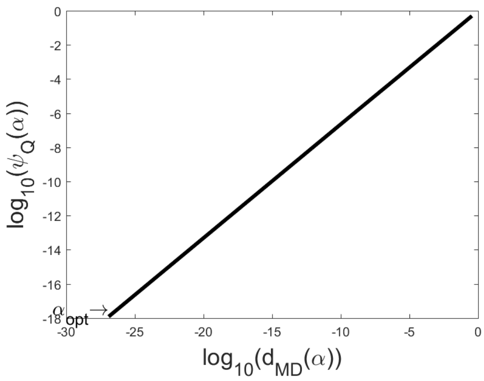

Figure 8. Q-curve in Groetsch2.

Figures 5–7 show examples of Q-curves and the triangle $T(k, r(k), l(k))$ with the largest area, the TA-rule chooses for the regularization parameter a corresponding minimizer. In some problems the function $\psi_Q(\alpha)$ may be monotonically increasing as for the problem Groetsch2 (Figure 8), then function $\psi_Q(\alpha)$ has only one local minimizer α_N. Then vertices $P(M_{l(k)})$ and $P(m_k)$ coincide and the area of the corresponding triangle is zero. Then this is the only triangle, the TA-rule chooses for the regularization parameter α_N. The results of the numerical experiments for test set 1 ($n = 100$) for the TA-rule and some other rules (see Section 2.2) are given in Tables 4 and 5. These results show that the TA-rule works well in all these test problems, the accuracy is comparable with δ-rules (see Table 3), but previous heuristic rules fail in some problems. Note that average of the error ratio increases for decreasing noise level. For example, for $\|f - f_*\| \in \{10^{-1}, 10^{-2}, 10^{-5}, 10^{-6}\}$ corresponding error ratios E were 1.47, 1.49, 1.78 and 2.08 respectively.

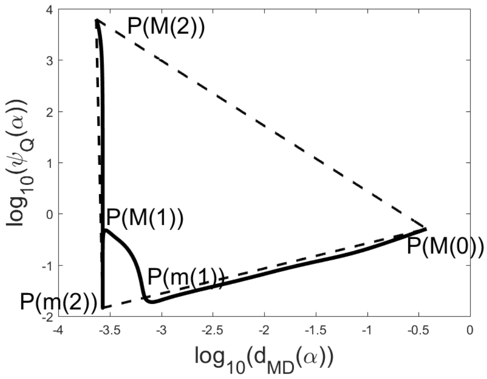

Figure 9. Q-curve, problem Heat, $n = 60$.

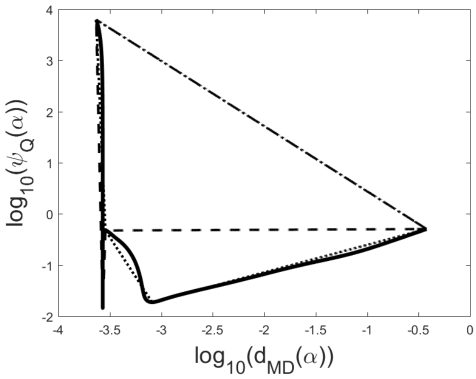

Figure 10. Q-curve, problem Heat, $n = 60$.

Let us comment on other heuristic rules. The accuracy of the quasi-optimality criterion is for many problems the same as for the TA-rule, but this rule fails in problem Heat. A characteristic feature of the problem Heat is that location of the eigenvalues in the interval $[\alpha_N, 1]$ is sparse and only some eigenvalues are smaller than α_N (see Table 1). The weighted quasioptimality criterion behaves in a similar way as the quasioptimality criterion, but is more accurate in problems where $\lambda_{\min} \leq \alpha_N$; if $\lambda_{\min} > \alpha_N$, the quasioptimality criterion is more accurate. The rule of Hanke–Raus may fail in test problems with large Λ and for other problems the error of the approximate solution is in most problems approximately two times larger than for parameter chosen by the quasi-optimality principle. The problem in this rule is that it chooses too large parameter compared with the optimal parameter. However, HR-rule is stable in the sense that the largest error ratio E2 is relatively small in all considered test problems. Reginska's rule may fail in many problems but it has the advantage that it works better than other previous rules if the noise level is large. The Reginska's rule did not fail in case $\|f - f_*\| \geq 10^{-3}$ and has average of error ratios of all problems $E = 2.24$ and $E = 2.80$ in cases $\|f - f_*\| = 10^{-1}$ and $\|f - f_*\| = 10^{-2}$, respectively. Advantage of the maximum curvature rule is the small percentage of failures compared with other previous rules.

Distribution of error ratios E in Table 5 show also that in test problems set 1 the TA-rule is the most accurate rule from all considered rules.

Table 4. Averages of error ratios E and failure % (in parenthesis) for heuristic rules.

Problem	TA Rule Mean E	TA Rule Max E	Quasiopt. Mean E	WQ Mean E	HR Mean E	Reginska Mean E	MCurv Mean E
Baart	1.51	14.57	1.54	1.43	2.58	1.32	4.75
Deriv2	1.18	1.27	2.01	2.26	2.28	3.67	(9.2%)
Foxgood	1.56	3.39	1.57	1.57	8.36	(10.8%)	5.95
Gravity	1.14	2.27	1.13	1.13	2.66	(0.8%)	2.04
Heat	1.26	1.34	(65.8%)	(66.8%)	1.64	(4.2%)	4.11
Ilaplace	1.24	2.34	1.24	1.22	1.94	1.66	2.99
Phillips	1.07	1.20	1.09	(3.3%)	2.27	(44.2%)	1.34
Shaw	1.42	8.96	1.43	1.41	2.34	1.80	4.64
Spikes	1.01	5.75	1.01	1.01	1.03	1.01	1.05
Wing	1.39	6.63	1.40	1.30	1.51	1.18	1.57
Baker	3.30	11.33	3.30	3.30	(0.8%)	(21.7%)	7.78
Ursell	2.87	31.06	3.54	2.35	4.71	1.86	7.54
Indramm	3.74	9.07	4.43	4.16	(2.5%)	(9.2%)	(15.8%)
Waswaz2	2.43	9.01	2.43	2.43	(65.8%)	2.33	(3.3%)
Groetsch1	1.14	4.56	1.14	1.12	1.61	1.26	1.52
Groetsch2	1.13	1.74	1.27	2.73	1.66	5.49	1.81
Total	1.71	31.06	>100	>100	50.5	43.8	8.17
Failure %	0%		4.11%	4.38%	4.32%	5.68%	1.77%
Max E2	2.61		>100	>100	2.63	>100	24.5

Table 5. Distribution of error ratios E in different rules.

Decile	TA Rule	Quasiopt.	WQ	HR	Reginska	MCurv	ME	MEe	DP
10	1.00	1.00	1.00	1.08	1.00	1.06	1.01	1.00	1.00
20	1.01	1.01	1.01	1.36	1.04	1.27	1.03	1.00	1.01
30	1.02	1.03	1.03	1.56	1.12	1.48	1.09	1.01	1.02
40	1.04	1.06	1.06	1.82	1.27	1.83	1.16	1.03	1.04
50	1.09	1.13	1.12	2.12	1.66	2.31	1.22	1.08	1.08
60	1.18	1.29	1.29	2.43	2.42	3.05	1.33	1.16	1.16
70	1.35	1.57	1.59	3.19	4.19	4.51	1.52	1.29	1.30
80	1.71	2.17	2.29	5.94	9.93	7.03	2.02	1.50	1.52
90	2.27	6.45	6.18	19.35	43.91	12.95	4.45	2.88	2.11

Note that the figure of the Q-curve enables to estimate the reliability of the chosen parameter. If the Q-curve has only one corner, then the chosen parameter is quasioptimal with small constant C, if $\lambda_{\min} < \alpha_N$, but in case $\lambda_{\min} \geq \alpha_N$ it is quasioptimal under assumption that the problem needs regularization.

6. Further Developments of the Area Rule

The TA-rule may fail for problems which do not need regularization, if function $\psi_Q(\alpha)$ is not monotonically increasing. In this case, the TA rule selects the parameter $\alpha \geq \lambda_{\min}$, but the parameter $\alpha < \lambda_{\min}$ would be better. For example, the TA-rule fails for matrix *Moler* in some cases. Let us consider now the question, in which cases the regularization parameter α_N is good. If function $\psi_Q(\alpha)$ is monotonically increasing then function $\psi_Q(\alpha)$ has only one local minimizer $m_1 = M_1 = \alpha_N$ and then for the parameter α_N we have the error estimate

$$\|u_{\alpha_N} - u_*\| \leq q^{-1}(1 + T(\alpha_N)) \min_{\alpha_N \leq \alpha \leq \alpha_0} e_1(\alpha),$$

where the value of $T(\alpha_N) = \max_{\alpha_j\in\Omega,\alpha_N\leq\alpha_j\leq\alpha_0} T\left(\alpha_N,\alpha_j\right) \leq c_q \ln(\frac{\alpha_0}{\alpha_N})$ (see Theorem 3) can be computed a posteriori and this value is the smaller the faster the function $\psi_{\mathrm{Q}}(\alpha)$ increases. We can take α_N for the regularization parameter also in the case if the condition

$$\frac{\psi_{\mathrm{Q}}(\alpha')}{\psi_{\mathrm{Q}}(\alpha)} \leq c_0 \qquad \forall\alpha,\alpha' \in \Omega, \quad \alpha_N \leq \alpha' < \alpha \leq \alpha_0 \tag{20}$$

holds while one can show similarly to the proof of Theorem 3 that $T(\alpha_N) \leq c_0 c_q \ln(\frac{\alpha_0}{\alpha_n})$ and the error of the regularized solution is small. For the problems which do not need regularization we can improve the performance of the TA-rule searching the proper local minimizer smaller or equal than $\alpha_{\mathrm{HQ}} := \max\{\alpha_{\mathrm{HR}},\alpha_{\mathrm{Q}}\}$, where α_{HR}, α_{Q} are the global minimizers of functions $\psi_{\mathrm{HR}}(\alpha)$ and $\psi_{\mathrm{Q}}(\alpha)$, respectively on the interval $[\max(\alpha_N,\lambda_{\min}),\alpha_0]$.

These ideas enable to formulate the following upgraded version of the TA-rule.

Triangle area rule 2 (TA-2-rule). We fix a constant $c_0, 1 \leq c_0 \leq 2$. If condition (20) holds, we choose the parameter α_N. Otherwise choose for the regularization parameter such local minimizer $m_k \leq \alpha_{\mathrm{HQ}}$ of the function $\psi_{\mathrm{Q}}(\alpha)$ for which the area of triangle $T(k,r(k),l(k))$ is largest.

Results of numerical experiments for the rule TA-2 with the discretization parameter $n = 100$ and problem sets 1–3 are given in Tables 6 and 7 (columns 2 and 3). The results show that the rule TA-2 works well in all considered testsets 1–3. However, the rule TA-2 may fail in some other problems which do not need regularization. Such example is the problem with matrix *moler* and solution $x_i = \sin(12\pi(i-1)/n)$, where the rule TA-2 fails if the noise level is below 10^{-4}; but in this case all other considered heuristic rules fail too.

The rules TA and TA-2 fail in problem Heat in some cases for the discretization parameter $n = 60$. Figures 9 and 10 show the form of the Q-curve in problem Heat with $n = 60$. The function $\psi_{\mathrm{Q}}(\alpha)$ has two local minimizers with corresponding points $P(m_1)$ and $P(m_2)$ on the Q-curve and 3 local maximum points $P(M_k), k = 0,1,2$. On Figure 9 the Rule TA-2 chooses the local minimizer corresponding to the point $P(m_2)$, but then the error ratios are large: $E = 20.3, E2 = 18.34$.

In the following we consider methods which work well also in this problem. Let $g[\alpha_1,\alpha_2](\alpha)$, $\alpha \in [\alpha_1,\alpha_2]$ be the parametric representation of the straight line segment connecting points $P(\alpha_1)$ ja $P(\alpha_2)$, thus

$$g[\alpha_1,\alpha_2](\alpha) = \tilde{\psi}_{\mathrm{Q}}(\alpha_1) + \beta(\tilde{d}_{\mathrm{MD}}(\alpha) - \tilde{d}_{\mathrm{MD}}(\alpha_1)), \quad \beta = \frac{\tilde{\psi}_{\mathrm{Q}}(\alpha_2) - \tilde{\psi}_{\mathrm{Q}}(\alpha_1)}{\tilde{d}_{\mathrm{MD}}(\alpha_2) - \tilde{d}_{\mathrm{MD}}(\alpha_1)}.$$

Let $g[\alpha_1,\alpha_2,...,\alpha_k](\alpha), k > 2$ be the parametric representation of the broken line connecting points $P(\alpha_1), P(\alpha_2),\ldots,P(\alpha_k)$, thus

$$g[\alpha_1,\alpha_2,...,\alpha_k](\alpha) = g[\alpha_j,\alpha_{j+1}](\alpha), \qquad \alpha_j \leq \alpha \leq \alpha_{j+1}, \quad 1 \leq j \leq k-1.$$

In the triangle rule certain points $P(M_{l(k)}), P(m_k), P(M_{r(k)})$ are connected by the broken line $t_1(\alpha) = g[M_{l(k)}, m_k, M_{r(k)}](\alpha)$ which approximates the error function $\tilde{e}_1(\alpha) = \log_{10}(e_1(\alpha))$ well, if m_k is the "right" local minimizer. By construction of the function $t_1(\alpha)$ we use only 3 points on the Q-curve. We will get a more stable rule if the form of the Q-curve has more influence to the construction of approximates to the error function $\tilde{e}_1(\alpha)$. Let $\{i(1),i(2),...,i(n1)\}$ and $\{j(1),j(2),...,j(n2)\}$ be the largest sets of the indices, satisfying the inequalities

$$k \leq i(1) < i(2) < ... < i(n1) \leq K, \quad \psi_{\mathrm{Q}}(M_{i(1)}) \leq \psi_{\mathrm{Q}}(M_{i(2)}) \leq ... \leq \psi_{\mathrm{Q}}(M_{i(n1)}),$$

$$k > j(1) > j(2) > ... > j(n2) \geq 0, \quad \psi_{\mathrm{Q}}(M_{j(1)}) \leq \psi_{\mathrm{Q}}(M_{j(2)}) \leq ... \leq \psi_{\mathrm{Q}}(M_{j(n2)}).$$

It is easy to see that $i(n1) = l(k)$ and $j(n2) = r(k)$. For approximating the error function $\tilde{e}_1(\alpha)$ we propose to connect points $P(M_{i(n1)}),...,P(M_{i(1)}),P(m_k),P(M_{j(1)}),...,P(M_{j(n2)})$ by the broken line $t_2(\alpha) = g[M_{i(n1)},...,M_{i(1)},m_k,M_{j(1)},...,M_{j(n2)}](\alpha)$ and to find for every m_k the area $S_2(k)$ of the polygon surrounded by the lines $T_2(\alpha) = \max\{t_2(\alpha), g[M_{i(n1)}, M_{j(n2)}](\alpha)\}$ and $t_2(\alpha)$. The second possibility is to approximate the error function $\tilde{e}_1(\alpha)$ by the curve $t_3(\alpha) = \max\{t_2(\alpha), \tilde{\psi}_Q(\alpha)\}$ and to find $S_3(k)$ as the area of the polygon surrounded by the broken lines $t_3(\alpha)$ and curve $T_3(\alpha) = \max\{T_2(\alpha), \tilde{\psi}_Q(\alpha)\}$. Note that functions $t_i(\alpha), i = 1, 2, 3$ are monotonically increasing if $\alpha > m_k$, and monotonically decreasing if $\alpha < m_k$.

Area rules 2 and 3. We fix a constant $c_0, 1 \leq c_0 \leq 2$. First we choose the local minimizer $m_k \leq \alpha_{HQ}$, for which the area $S_i(k),\ i \in \{2, 3\}$ is largest. We take for the regularization parameter the smallest $m_{k_0} \leq m_k$, satisfying condition (compare with (20))

$$\frac{\psi_Q(\alpha')}{\psi_Q(\alpha)} \leq c_0 \qquad \forall \alpha, \alpha' \in \Omega, \quad m_{k_0} \leq \alpha' < \alpha \leq m_k.$$

Let us consider Figure 9. The reason for the failure of the triangle rule is, that for the local minimizer $m(2)$ the broken line $g[M(0), m(2), M(2)](\alpha)$ do not approximate well the function $\tilde{e}_1(\alpha)$, while point $M(1)$ is located above the interval $[m(2), M(0)]$. Here the function $\tilde{e}_1(\alpha)$ is better approximated by the broken line $g[M(0), M(1), m(2), M(2)](\alpha)$, see Figure 10. For the local minimizer $m(1)$, we approximate function $\tilde{e}_1(\alpha)$ by the broken line $g[M(0), m(1), M(1), M(2)](\alpha)$ and due to the inequality $S_2(1) > S_2(2)$ the area rule 2 chooses m_1 for the regularization parameter, then $E = 1.05$.

The area rules 2 and 3 work in problem Heat well for every n and all $\alpha_N = 10^{-k}, 12 \leq k \leq 24$, but in some other problems the accuracy of the area rules 2 and 3 (see columns 4–7 of Tables 6 and 7) is slightly worse than for the rule TA-2. The advantage of the area rule 3, as compared to the area rule 2, is to be highlighted in problem Heat if all noise of the right hand side is placed on one eigenelement (then we use the condition $m_k \leq \alpha_{HQ}$ only in case $\lambda_{\min} > \alpha_N$). Then the area rule 3 did not fail if $n \geq 80$ and $\alpha_N \leq 10^{-20}$. So we can say, the more precisely we take into account the form of the Q-curve in construction of the approximating function for the error function $\tilde{e}_1(\alpha)$, the more stable is the rule.

Table 6. Averages and maximums of error ratios E in case of area rules, problem set 1.

Problem	TA-2 Rule		Area Rule 2		Area Rule 3		Combined Area Rule	
	Aver E	Max E	Aver E	Max E	Aver E	Max E	Aver E	Max E
Baart	1.51	5.18	1.58	2.91	1.59	2.91	1.53	5.18
Deriv2	1.12	1.42	1.12	1.42	1.12	1.42	1.12	1.42
Foxgood	1.57	6.69	1.53	6.19	1.53	6.19	1.57	6.69
Gravity	1.17	4.12	1.21	6.10	1.21	6.10	1.17	4.12
Heat	1.12	2.36	1.12	2.36	1.12	2.36	1.12	2.36
Ilaplace	1.24	2.68	1.22	2.68	1.22	2.68	1.24	2.68
Phillips	1.07	1.72	1.06	1.72	1.06	1.72	1.07	1.72
Shaw	1.42	3.72	1.47	3.64	1.47	3.64	1.42	3.72
Spikes	1.01	1.05	1.01	1.02	1.01	1.02	1.01	1.05
Wing	1.39	1.86	1.44	1.86	1.44	1.86	1.39	1.86
Baker	3.30	45.29	2.67	22.67	2.67	22.67	2.91	33.12
Ursell	2.87	16.78	4.55	27.92	4.55	27.92	3.12	16.78
Indramm	3.74	25.67	9.50	83.20	10.76	83.20	3.87	25.67
Waswaz2	2.43	9.01	2.43	9.01	2.43	9.01	2.43	9.01
Groetsch1	1.14	2.12	1.15	2.12	1.15	2.12	1.14	2.12
Groetsch2	1.52	3.84	1.52	3.84	1.52	3.84	1.52	3.84
Total	1.73	45.29	2.16	83.20	2.22	83.20	1.73	33.12

Based on the above rules it is possible to formulate a combined rule, which chooses the parameter according to the rule TA-2 or area rule 3 in dependence of certain condition.

Area rule 4 (Combined area rule). Fix a constant $c_0, 1 \leq c_0 \leq 2, b \geq 0$. Let the local minimizer m_k be chosen by the rule TA-2. If

$$\max_{m_k \leq \alpha \leq M_{r(k)}} \frac{\tilde{\psi}_Q(\alpha)}{g[m_k, M_{r(k)}](\alpha)} \leq b,$$

we take m_k for the regularization parameter, otherwise we choose the regularization parameter by the area rule 3.

Note that the combined rule coincides with the rule TA-2, if $b = 0$ and with the area rule 3, if $b = \infty$. Experiments of combined rule with $c_0 = 2, b = 1$ (columns 8 and 9 in Tables 6 and 7) show that the accuracy of this rule is almost the same as in the triangle rule, but unlike the TA-2 rule, it works well also in problem Heat for all n and α_N. Although, in some cases, in test set 3 the error ratio $E > 100$ for rule 4, the high qualification of the rule is characterized by the fact, that over all problems sets 1-3 the largest error ratio E1 was 16.91 (5.06 for set 1) and the largest error ratio E2 was 4.67 (2.62 for set 1). Numerical experiments show that it is reasonable to use parameter $b \in (0.8, 1.2]$. We studied the behavior of the area rules for different $\alpha_N = 10^{-k}, 12 \leq k \leq 24$. The results were similar to the results of Tables 6 and 7, but for smaller α_N the error ratios were 2–3% smaller than for $\alpha_N = 10^{-18}$ and for larger α_N the error ratios were about 5% larger than in Tables 6 and 7.

Table 7. Averages and maximums of error ratios E in proposed rules, problem sets 2 and 3.

Problem	TA-2 Rule		Area Rule 2		Area Rule 3		Combined Area Rule	
	Aver E	Max E	Aver E	Max E	Aver E	Max E	Aver E	Max E
Gauss	1.24	5.05	1.26	6.56	1.26	6.56	1.24	5.05
Hilbert	1.46	7.25	1.83	21.22	1.81	21.22	1.46	7.25
Lotkin	1.47	11.17	1.91	18.66	1.88	11.17	1.47	11.17
Moler	1.51	7.35	1.43	7.35	1.43	7.35	1.51	7.35
Prolate	1.57	15.96	1.82	20.64	1.77	15.96	1.58	15.96
Pascal	1.04	1.13	1.06	1.18	1.06	1.18	1.05	1.18
Set 2	1.38	15.96	1.55	21.22	1.53	21.22	1.39	15.96
Set 3	1.85	136.6	2.81	188.1	2.77	188.1	2.02	153.5

The Table 2 gives results of the numerical experiments in the case of smooth solution, $p = 2$. We see that combined rule worked well also in this case, no failure.

Remark 5. *It is possible to modify the Q-curve. We may use the function $\psi_{QD}(\alpha)$ instead of function $\psi_Q(\alpha)$ and find proper local minimizer of the function $\psi_{QD}(\alpha)$. Unlike the quasi-optimality criterion the use of function $\psi_{QD}(\alpha)$ in the Q-curve and in the area rule does not increase the amount of calculations, while approximation $u_{2,\alpha}$ is needed also in computation of $d_{MD}(\alpha)$. We can use in these rules the function $d_{ME}(\alpha)$ instead of $d_{MD}(\alpha)$, it increases the accuracy in some problems, but the average accuracy of the rules is almost the same. In case of nonsmooth solutions we can modify the Q-curve method and area rule, using the function $d_D(\alpha)$ instead of $d_{MD}(\alpha)$. In this case, we get even better results for $p = 0$ but for $p = 2$, the error ratio E is on average 2 times higher.*

Note that if the solution is smooth, then L-curve rule and Reginska's rule often fail, but replacing in these rules the function $Au_\alpha - f$ by the function $B_\alpha(Au_\alpha - f)$ gives often better results.

Remark 6. *If the solution is smooth, then using α-s from (8) much better approximate solution than single Tikhonov approximation may be get using the linear combinations of Tikhonov approximations, see [40].*

In the case of a heuristic parameter choice, it is also possible to use the a posteriori estimates of the approximate solution, which, in many tasks, allows to confirm the reliability of the parameter choice. Let α_H be the regularization parameter from some heuristic rule and α_* be the local minimizer of the function $e_1(\alpha)$ on the set Ω. Then in the case $\alpha_* \geq \alpha_H$ the error estimate

$$\|u_{\alpha_H} - u_*\| \leq (1 + T(\alpha_H, \alpha_*))\, e_1(\alpha_*) \leq q^{-1}\,(1 + T_1(\alpha_H)) \min_{\alpha} e_1(\alpha) \tag{21}$$

holds where $T_1(\alpha_H) = \max_{\alpha \geq \alpha_H, \alpha \in \Omega} T(\alpha_H, \alpha)$. Using the last estimate, we can prove similarly to the Theorem 7 that if α_N is so small that $d_{\mathrm{MD}}(\alpha_N) \leq (1 + \epsilon)\, \|f - f_*\|$, then

$$\|u_{\alpha_H} - u_*\| \leq \max\{q^{-1}(1 + T_1(\alpha_H)) \min_{\alpha \geq 0} e_1(\alpha), C_2(b, \epsilon) \min_{\alpha \geq 0} e_2(\alpha, \|f - f_*\|)\},$$

where $b = d_{\mathrm{MD}}(\alpha_H)/d_{\mathrm{MD}}(\alpha_N)$. If values $T_1(\alpha_H)$ and b, what we find a posteriori, are small (for example $b \leq 2$ and $T_1(\alpha_H) \leq 9$), then this estimate allows to argue that the error of the approximate solution for this parameter is not much larger than the minimal error. The conditions $b \leq 2$, $T_1(\alpha_H) \leq 9$ were satisfied in set 1 of test problems in the combined rule for 73% of cases and inequalities $b \leq 2$, $T_1(\alpha_H) \leq 4$ for 61% of cases. The reason of failure of heuristic rule is typically that the chosen parameter is too small. To check this, we can use the error estimate (21). If $T_1(\alpha_H)$ is relatively small (for example $T_1(\alpha_H) \leq 9$), then the estimate (21) allows to argue that the regularization parameter is not chosen too small. In set 1 of test problems the conditions $T_1(\alpha_H) \leq 9$ and $T_1(\alpha_H) \leq 4$ were satisfied in 97% and in 82% of cases, respectively.

7. Conclusions

We finish the paper with the following conclusion. For the heuristic choice of the regularization parameter we recommend to choose the parameter from the set of local minimizers of the function $\psi_{\mathrm{Q}}(\alpha)$ or the function $\psi_{\mathrm{QD}}(\alpha)$. For choice of the parameter from the local minimizers we proposed the Q-curve method and different area rules. The proposed rules gave much better results than previous heuristic rules on extensive set of test problems. Area rules fail in very few cases in comparison with previous rules, and the accuracy of these rules is comparable even with the δ-rules if the exact noise level is known. In addition, we also provided a posteriori error estimates of the approximate solution, which allows to check the reliability of parameter chosen heuristically.

Author Contributions: Writing—original draft preparation, T.R.; writing—review and editing, U.H. Both authors have read and agreed to the published version of the manuscript.

Funding: The authors are supported by "Personal research funding: Team grant" project PRG864 of the Estonian Ministry of Education and Research.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Engl, H.W.; Hanke, M.; Neubauer, A. *Regularization of Inverse Problems, Volume 375 of Mathematics and Its Applications*; Kluwer: Dordrecht, The Netherlands, 1996.
2. Vainikko, G.M.; Veretennikov, A.Y. *Iteration Procedures in Ill- Posed Problems*; Nauka: Moscow, Russia, 1986. (In Russian)
3. Bauer, F.; Kindermann, S. The quasi-optimality criterion for classical inverse problems. *Inverse Probl.* **2008**, *24*, 035002.
4. Hämarik, U.; Palm, R.; Raus, T. On minimization strategies for choice of the regularization parameter in ill-posed problems. *Numer. Funct. Anal. Optim.* **2009**, *30*, 924–950.
5. Kindermann, S. Convergence analysis of minimization-based noise level-free parameter choice rules for linear ill-posed problems. *Electron. Trans. Numer. Anal.* **2011**, *38*, 233–257.
6. Kindermann, S. Discretization independent convergence rates for noise level-free parameter choice rules for the regularization of ill-conditioned problems. *Electron. Trans. Numer. Anal.* **2013**, *40*, 58–81.
7. Kindermann, S.; Neubauer, A. On the convergence of the quasioptimality criterion for (iterated) Tikhonov regularization. *Inverse Probl. Imaging* **2008**, *2*, 291–299.
8. Neubauer, A. The convergence of a new heuristic parameter selection criterion for general regularization methods. *Inverse Probl.* **2008**, *24*, 055005.

9. Tikhonov, A.N.; Glasko, V.B.; Kriksin, Y. On the question of quasioptimal choice of a regularized approximation. *Sov. Math. Dokl.* **1979**, *20*, 1036–1040.
10. Hansen, P.C. Analysis of discrete ill-posed problems by means of the L-curve. *SIAM Rev.* **1992**, *34*, 561–580.
11. Hansen, P.C. Regularization tools: A Matlab package for analysis and solution of discrete ill-posed problems. *Numer. Algorithms* **1994**, *6*, 1–35.
12. Golub, G.H, Heath, M.; Wahba, G. Generalized cross-validation as a method for choosing a good ridge parameter. *Technometrics* **1979**, *21*, 215– 223.
13. Hanke, M.; Raus, T. A general heuristic for choosing the regularization parameter in ill-posed problems. *SIAM J. Sci. Comput.* **1996**, *17*, 956–972.
14. Reginska, T.A. A regularization parameter in discrete ill-posed problems. *SIAM J. Sci. Comput.* **1986**, *17*, 740–749.
15. Gockenbach, M.S.; Gorgin, E. On the convergence of a heuristic parameter choice rule for Tikhonov regularization. *SIAM J. Sci. Comput.* **2018**, *40*, A2694–A2719.
16. Hämarik, U.; Kangro, U.; Kindermann, S.; Raik, K. Semi-heuristic parameter choice rules for Tikhonov regularisation with operator perturbations. *J. Inverse Ill Posed Probl.* **2019**, *27*, 117–131.
17. Hochstenbach, M.E, Reichel, L.; Rodriguez, G. Regularization parameter determination for discrete ill-posed problems *J. Comput. Appl. Math.* **2015**, *273*, 132–149.
18. Palm, R. Numerical Comparison of Regularization Algorithms for Solving ill-Posed Problems. Ph.D. Thesis, University of Tartu, Tartu, Estonia, 2010.
19. Bauer, F.; Lukas, M.A. Comparing parameter choice methods for regularization of ill-posed problems. *Math. Comput. Simul.* **2011**, *81*, 1795–1841.
20. Bakushinskii, A.B. Remarks on choosing a regularization parameter using the quasi-optimality and ratio criterion. *Comp. Math. Math. Phys.* **1984**, *24*, 181–182.
21. Raus, T.; Hämarik, U. Heuristic parameter choice in Tikhonov method from minimizers of the quasi-optimality function. In *New Trends in Parameter Identification for Mathematical Models*; Hofmann, B., Leitao, A., Zubelli, J., Eds.; Birkhäuser: Basel, Switzerland, 2018; pp. 227–244.
22. Brezinski B, Rodriguez, G.; Seatzu, S. Error estimates for linear systems with applications to regularization. *Numer. Algor.* **2008**, *49*, 85–104.
23. Baker, C.T.H.; Fox, L.; Mayers, D.F.; Wright, K. Numerical solution of Fredholm integral equations of the first kind. *Comput. J.* **1964**, *7*, 141–148.
24. Groetsch, C.W. Integral equations of the first kind, integral equations, and regularization: A crash course. *J. Phys. Conf. Ser.* **2007**, *73*, 012001.
25. Indratno, S.W.; Ramm, A.G. An iterative method for solving Fredholm integral equations of the first kind. *Int. J. Comput. Sci. And Mathematics* **2009**, *2*, 354–379.
26. Wazwaz, A.-M. The regularization method for Fredholm integral equations of the first kind. *Comput. Math. Appl.* **2011**, *61*, 2981–2986.
27. Morozov, V.A. On the solution of functional equations by the method of regularization. *Soviet Math. Dokl.* **1966**, *7*, 414–417.
28. Gfrerer, H. An a posteriori parameter choice for ordinary and iterated Tikhonov regularization of ill-posed problems leading to optimal convergence rates. *Math. Comp.* **1987**, *49*, 507–522.
29. Raus, T. On the discrepancy principle for solution of ill-posed problems with non-selfadjoint operators. *Acta Comment Univ. Tartu.* **1985**, *715*, 12–20. (In Russian)
30. Hämarik, U.; Kangro, U.; Palm, R.; Raus, T.; Tautenhahn, U. Monotonicity of error of regularized solution and its use for parameter choice. *Inverse Probl. Sci. Eng.* **2014**, *22*, 10–30.
31. Tautenhahn, U.; Hämarik, U. The use of monotonicity for choosing the regularization parameter in ill-posed problems. *Inverse Probl.* **1999**, *15*, 1487–1505.
32. Hämarik, U.; Palm, R.; Raus, T. Comparison of parameter choices in regularization algorithms in case of different information about noise level. *Calcolo* **2011**, *48*, 47–59.
33. Hämarik, U.; Palm, R.; Raus, T. A family of rules for parameter choice in Tikhonov regularization of ill-posed problems with inexact noise level. *J. Comp. Appl. Math.* **2012**, *36*, 221–233.
34. Raus, T.; Hämarik, U. New rule for choice of the regularization parameter in (iterated) Tikhonov method. *Math. Model. Anal.* **2009**, *14*, 187–198.

35. Raus, T.; Hämarik, U. On numerical realization of quasioptimal parameter choices in (iterated) Tikhonov and Lavrentiev regularization. *Math. Model. Anal.* **2009**, *14*, 99–108.
36. Raus, T. About regularization parameter choice in case of approximately given error bounds of data. *Acta Comment Univ. Tartu.* **1992**, *937*, 77–89.
37. Raus, T.; Hämarik, U. On the quasioptimal regularization parameter choices for solving ill-posed problems. *J. Inverse Ill Posed Probl.* **2007**, *15*, 419–439.
38. Calvetti, D.; Reichel, L.; Shuib, A. L-curve and curvature bounds for Tikhonov regularization. *Numer. Algorithms* **2004**, *35*, 301–314.
39. Hansen, P.C. *Rank-Deficient and Discrete Ill-Posed Problems* SIAM: Philadelphia, PA, USA, 1998.
40. Hämarik, U.; Palm, R.; Raus, T. Extrapolation of Tikhonov regularization method. *Math. Model. Anal.* **2010**, *15*, 55–68.