

# Regularized Fixed-Point Iterations for Nonlinear Inverse Problems

S.S. Pereverzyev<sup>\*</sup>, R. Pinnau<sup>†</sup> and N. Siedow<sup>‡</sup>

August, 2005

## Abstract

In this paper we introduce a derivative-free, iterative method for solving nonlinear ill-posed problems  $Fx = y$ , where instead of  $y$  noisy data  $y_\delta$  with  $\|y - y_\delta\| \leq \delta$  are given and  $F : D(F) \subseteq X \rightarrow Y$  is a nonlinear operator between Hilbert spaces  $X$  and  $Y$ . This method is defined by splitting the operator  $F$  into a linear part  $A$  and a nonlinear part  $G$ , such that  $F = A + G$ . Then iterations are organized as  $Au_{k+1} = y_\delta - Gu_k$ . In the context of ill-posed problems we consider the situation when  $A$  does not have a bounded inverse, thus each iteration needs to be regularized. Under some conditions on the operators  $A$  and  $G$  we study the behavior of the iteration error. We obtain its stability with respect to the iteration number  $k$  as well as the optimal convergence rate with respect to the noise level  $\delta$ , provided that the solution satisfies a generalized source condition. As an example, we consider an inverse problem of initial temperature reconstruction for a nonlinear heat equation, where the nonlinearity appears due to radiation effects. The obtained iteration error in the numerical results has the theoretically expected behavior. The theoretical assumptions are illustrated by a computational experiment.

**AMS MSC:** 65J15, 65J20, 80A23.

**Keywords:** nonlinear inverse problem, regularization, a posteriori regularization parameter choice, derivative-free iterative method, generalized source condition, order optimal convergence, nonlinear heat equation, initial temperature, heat radiation, computational experiment.

---

<sup>\*</sup> *Fraunhofer-Institut für Techno- und Wirtschaftsmathematik, Gottlieb-Daimler-Straße, D-67663 Kaiserslautern, Germany; e-mail: pereverz@itwm.fhg.de*

<sup>†</sup> *Fachbereich Mathematik, Technische Universität Kaiserslautern, D-67663 Kaiserslautern, Germany; e-mail: pinnau@mathematik.uni-kl.de*

<sup>‡</sup> *Fraunhofer-Institut für Techno- und Wirtschaftsmathematik, Gottlieb-Daimler-Straße, D-67663 Kaiserslautern, Germany; e-mail: siedow@itwm.fhg.de*

# 1 Introduction

Let us consider a nonlinear inverse problem represented by the operator equation

$$Fu = y, \quad (1)$$

with a nonlinear operator  $F : D(F) \subseteq X \rightarrow Y$  acting between some Hilbert spaces  $X$  and  $Y$ . The inner product and the corresponding norm on each of these spaces will be denoted by  $(\cdot, \cdot)$  and  $\|\cdot\|$ , respectively. It will be always clear from the context which space is considered.

We assume that (1) possesses a solution  $\hat{u}$ , and this solution is unique at least in some ball around  $\hat{u}$ . We are interested in the case when (1) is ill-posed, i.e., the solution of (1) does not depend continuously on the data  $y$ . In particular, given noisy data  $y_\delta$ , such that  $\|y - y_\delta\| \leq \delta$ , the equation

$$Fu = y_\delta \quad (2)$$

can have no solution, and even if the solution exists, its distance to  $\hat{u}$  can be arbitrary large. Thus, special methods, so-called *regularization methods*, are required for stabilizing (2) [8, 14, 10, 26].

There are basically two approaches to the regularization of (2).

**Approach 1 (“regularization first”):** Instead of the ill-posed nonlinear equation (2) one considers some family of well-posed nonlinear equations

$$F_\alpha u = y_\delta, \quad (3)$$

with solutions  $\{u^\alpha\}$  depending on the regularization parameter  $\alpha$ , and searches for such  $\alpha = \alpha(\delta)$  that the norm  $\|\hat{u} - u^\alpha\|$  takes the (order-)optimal value with respect to the noise level  $\delta$ . Realizations of this approach are Tikhonov regularization [32, 9, 20, 28], Lavrentiev regularization [31], and the general regularization scheme introduced by Tautenhahn [29].

**Approach 2 (“linearization first”):** Here the operator  $F$  is approximated by some linear operator  $L_v$

$$Fu \approx L_v u + b_v, \quad (4)$$

which can depend on the point  $v \in X$  around which the linearization is made. One of the frequently used linearization schemes involves the Fréchet derivative of  $F$ . In this case  $L_v = F'_v := F'(v)$  and  $b_v = Fv - F'_v v$ . Having (4), one associates with (2) the following iterative procedure

$$L_{u_k} u_{k+1} = y_\delta - b_{u_k}, \quad k = 0, 1, \dots \quad (5)$$

Frequently, the linearization of nonlinear ill-posed equations also leads to ill-posed equations. That’s why each iteration (5) needs to be regularized. Thus, instead of (5) one has to consider

$$L_{u_k}^\alpha u_{k+1} = y_\delta - b_{u_k}, \quad k = 0, 1, \dots, \quad (6)$$

where  $L_{u_k}^\alpha$  is some regularized approximation of  $L_{u_k}$ . But in (5) the noise of the right hand side is uncontrollable, because in addition to the data noise we have the linearization error that is unknown. Thus, within each iteration (6) it is impossible to choose the regularization parameter  $\alpha$  in such a way that the solutions of (6) would converge to the solution of the linearized equation as  $\delta \rightarrow 0$ , because for such a choice the knowledge of the noise level is required [3]. The way out of this situation is known as *iterative regularization*. It was proposed by Bakushinskii [4] and studied in [5, 13, 12, 7]. The idea is to choose a sequence of regularization parameters  $\{\alpha_k\}$  that satisfies

$$1 \leq \frac{\alpha_k}{\alpha_{k+1}} \leq C, \quad \lim_{k \rightarrow \infty} \alpha_k = 0,$$

with some constant  $C > 1$ . Then the iteration procedure (6) runs with this sequence and the iteration number  $k$  is used as a regularization parameter.

In this paper we propose another approach for regularizing (2). We show, that for linearizations satisfying some conditions it is possible to choose the regularization parameter within each iteration in such a way that it will guarantee stability of the iterative process. Namely, the *iteration error*  $\|\hat{u} - u_k\|$  will decrease until some level and then remains below this level. Such a stable behavior

is unusual for iterative processes used for ill-posed nonlinear equations, where one has to stop iterations somewhere to avoid error explosion.

We consider a linearization of the operator  $F$  achieved by its splitting into a linear part  $A$  and a nonlinear part  $G$ , i.e.,

$$Fu = Au + Gu. \quad (7)$$

This splitting suggests a fixed-point iteration for solving (2):

$$Au_{k+1} = y_\delta - Gu_k. \quad (8)$$

We are especially interested in the case when the operator  $A$  does not have a bounded inverse. Thus, each fixed-point iteration (8) needs to be regularized. Within the framework of the general regularization scheme [8, p. 71] a regularization method is generated by some operator valued function  $\alpha \rightarrow g_\alpha(A^*A)$ , and the regularized approximate solution of (8) is given by choosing some  $\alpha = \alpha(\delta, k)$  and letting

$$u_{k+1} = g_\alpha(A^*A)A^*(y_\delta - Gu_k). \quad (9)$$

The paper is organized as follows. In Section 2 we derive an estimate for the iteration error  $\|\hat{u} - u_{k+1}\|$ . The similarity of this estimate to the regularization error for linear problems allows the application of the recently proposed adaptive strategy [18, 24], which is done in Section 3. Under the general smoothness assumptions for the solution  $\hat{u}$ , we are able to show that the iterative process (9) gives the order-optimal convergence rate with respect to the noise level  $\delta$ . We show also that the error behavior is stable with respect to the iteration number  $k$ . The a-priori choice of the regularization parameter is discussed in Section 4. This choice allows to obtain similar results under more relaxed assumptions on the nonlinearity. In Section 5 we consider an inverse problem of the initial temperature reconstruction for a nonlinear heat equation where the splitting (7) naturally arises. We consider specifically the heat equation including radiation effects. In this case the application of standard methods involving Fréchet derivatives is quite complicated, thus the proposed splitting is an attractive alternative. Numerical results show the theoretically predicted error behavior. The assumptions on the nonlinearity used in the theoretical considerations are illustrated by a computational experiment. We give some final remarks in Section 6.

## 2 Structure of the Iteration Error

In (7) we restrict ourselves to the case of the compact operator  $A$  with infinite dimensional range. Then,  $A$  admits a singular value decomposition

$$Ax = \sum_{k=1}^{\infty} s_k(x, v_k)w_k,$$

here  $\{v_k\} \subset X$ ,  $\{w_k\} \subset Y$  are some orthonormal systems, and a sequence  $\{s_k\}$  is such that it holds  $s_k > 0$ ,  $s_k \rightarrow 0$  as  $k \rightarrow \infty$ . Then,

$$g_\alpha(A^*A)x = \sum_{k=1}^{\infty} g_\alpha(s_k^2)(x, v_k)v_k.$$

Set  $a = \|A\|^2$ . Following [8, p. 72-74], we assume that for the family  $\{g_\alpha\}$  there are positive constants  $\gamma_1, \gamma_2, \gamma_3$  and  $\mu$  such that for all  $t \in [0, a]$  it holds that

$$|t g_\alpha(t)| \leq \gamma_1, \quad (R1)$$

$$|\sqrt{t} g_\alpha(t)| \leq \frac{\gamma_2}{\sqrt{\alpha}}, \quad (R2)$$

$$|1 - t g_\alpha(t)| t^\mu \leq \gamma_3 \alpha^\mu. \quad (R3)$$

**Example.** For the famous Tikhonov-Phillips regularization the function  $g_\alpha(t)$  has the following form

$$g_\alpha(t) = \frac{1}{\alpha + t}.$$

Then (R1)-(R3) are satisfied with  $\gamma_1 = 1$ ,  $\gamma_2 = 1/2$ ,  $\gamma_3 = 1$ ,  $\mu = 1$ .

Observe that the error  $\|\hat{u} - u_{k+1}\|$  of the iteration process (9) can be represented as

$$\begin{aligned}\|\hat{u} - u_{k+1}\| &= \|\hat{u} - g_\alpha(A^*A)A^*(y_\delta - Gu_k)\| \\ &\leq \|(I - g_\alpha(A^*A)A^*A)\hat{u}\| + \\ &\quad \|g_\alpha(A^*A)A^*(y - y_\delta)\| + \\ &\quad \|g_\alpha(A^*A)A^*(G\hat{u} - Gu_k)\|.\end{aligned}\tag{10}$$

Due to the property (R2), the so-called *stability term*  $\|g_\alpha(A^*A)A^*(y - y_\delta)\|$  can be estimated as

$$\|g_\alpha(A^*A)A^*(y - y_\delta)\| \leq \frac{\gamma_2}{\sqrt{\alpha}}\delta.\tag{11}$$

To simplify the notations we introduce  $\lambda(\alpha) = \frac{\sqrt{\alpha}}{\gamma_2}$ . From [27] it follows that a non-trivial estimate of the *approximation error*  $\|(I - g_\alpha(A^*A)A^*A)\hat{u}\|$  can only be obtained under an additional assumption regarding the smoothness of the solution  $\hat{u}$ . The most general form of such an assumption is known as *generalized source condition* and means that the solution  $\hat{u}$  belongs to the set

$$\mathcal{A}_\varphi = \{x \in X \mid x = \varphi(A^*A)v, \|v\| \leq 1\},\tag{12}$$

where  $\varphi : [0, a] \rightarrow \mathbb{R}_+$  is the so-called *index function*, which we assume to be continuous, increasing and satisfying  $\varphi(0) = 0$ . We also assume that  $\varphi$  satisfies the so-called  $\Delta_2$  condition, i.e., there is some constant  $C$  such that

$$\varphi(2t) \leq C \cdot \varphi(t)\tag{13}$$

for all  $t \in [0, a/2]$ .

**Example.** For simpler notation we use the asymptotic binary relation  $a(t) \asymp b(t)$ , i.e., there exist some constants  $\beta_1, \beta_2 > 0$  such that

$$\beta_1 b(t) \leq a(t) \leq \beta_2 b(t).$$

Index functions  $\varphi(t) \asymp t^\nu$ ,  $\nu > 0$  are frequently used. In this case one says that the problem of recovering the solution, which satisfies corresponding source condition (12), is *moderately ill-posed*. Other frequently used index functions are  $\varphi(t) \asymp \log^{-\nu} \frac{1}{t}$ ,  $\nu > 0$ , which give rise to *severely (exponentially) ill-posed problem*.

One more assumption is borrowed from [18] and describes the interplay of the index function  $\varphi$  and the function  $g_\alpha$  generating the regularization method. We assume that there is a constant  $\gamma_\mu > 0$  such that

$$\gamma_\mu \cdot \frac{\alpha^\mu}{\varphi(\alpha)} \leq \inf_{t \in [\alpha, a]} \frac{t^\mu}{\varphi(t)}, \quad \alpha \in (0, a],\tag{14}$$

where  $\mu$  is the constant from (R3).

**Example 1.** Condition (14) is satisfied for  $g_\alpha$  taken as in the Tikhonov-Phillips regularization and  $\varphi(t) \asymp \log^{-\nu} \frac{1}{t}$ . This means, in particular, that for severely ill-posed problems, Tikhonov-Phillips regularization gives order-optimal accuracy, which was already pointed out in [23].

As it has been proven in [18] for any  $\varphi$  meeting (14) the following inequality holds true

$$\sup_{t \in (0, a]} |1 - t g_\alpha(t)| \varphi(t) \leq \frac{\gamma_3}{\gamma_\mu} \cdot \varphi(\alpha), \quad \alpha \in (0, a].$$

To simplify the notation we will deal with  $\psi(\alpha) = \frac{\gamma_3}{\gamma_\mu} \cdot \varphi(\alpha)$ . Then, the approximation error can be estimated as

$$\|(I - g_\alpha(A^*A)A^*A)\hat{u}\| \leq \sup_{t \in (0, a]} |1 - t g_\alpha(t)| \varphi(t) \leq \psi(\alpha).\tag{15}$$

The third term in (10) depends on the interaction of the linear and nonlinear components of the splitting (7). To estimate it we introduce the following assumption:

**Assumption.** For any  $u \in B_r(\hat{u}) = \{x \in X \mid \|x - \hat{u}\| \leq r\}$  there is  $v(u) \in X$  such that

$$G\hat{u} - Gu = Av(u),\tag{A1}$$

$$\|v(u)\| \leq \rho \|\hat{u} - u\|,\tag{A2}$$

where  $\rho > 0$  is some constant.

The essence of the imposed assumption is that  $G\hat{u} - Gu$  lies always in the range of the operator  $A$ . This ensures well-posedness of the iteration steps in the procedure (8) applied for the noise-free data  $y$ .

This assumption leads to the estimate

$$\|g_\alpha(A^*A)A^*(G\hat{u} - Gu_k)\| \leq \gamma_1\rho\|\hat{u} - u_k\|. \quad (16)$$

Combining (11), (15) and (16) we obtain

$$\|\hat{u} - u_{k+1}\| \leq \psi(\alpha) + \frac{\delta}{\lambda(\alpha)} + \gamma_1\rho\|\hat{u} - u_k\|. \quad (17)$$

This estimate will be used for analysing the behavior of the iteration error. It is clear that due to the data noise,  $\|\hat{u} - u_{k+1}\|$  can not go to zero. In the case when our initial guess is the solution  $\hat{u}$  itself the data noise in the equation

$$Au = y_\delta - G\hat{u}$$

does not allow the exact reconstruction. Under rather general conditions on the operator  $A$  and the index function  $\varphi$ , [18, Theorem 1] gives the best possible order of accuracy  $b(\delta) = \varphi(\theta^{-1}(\delta))$ , that can be guaranteed for the recovery of the solution  $\hat{u} \in \mathcal{A}_\varphi$  from this noisy equation. Here  $\theta(t) = \sqrt{t}\varphi(t)$ .

**Example.** For moderately ill-posed problems the best possible order of accuracy is  $b(\delta) \asymp \delta^{\frac{\nu}{\nu+1/2}}$ , as it can be easily checked. For severely ill-posed problems  $b(\delta) \asymp \log^{-\nu} \frac{1}{\delta}$  and we refer to [16, 30] for details.

In the following it will be shown that the order  $b(\delta)$  can be reached by the regularized fixed-point iteration (9) with properly chosen regularization parameters.

### 3 Adaptive Choice of the Regularization Parameters

In view of estimate (17) it would be reasonable to choose the regularization parameters  $\alpha_{\text{opt}}^{(k+1)}$  such that

$$\psi\left(\alpha_{\text{opt}}^{(k+1)}\right) + \gamma_1\rho\|\hat{u} - u_k\| = \frac{\delta}{\lambda\left(\alpha_{\text{opt}}^{(k+1)}\right)}. \quad (18)$$

Let us denote by  $\alpha_{\text{opt}}$  the regularization parameter satisfying

$$\psi(\alpha_{\text{opt}}) = \frac{\delta}{\lambda(\alpha_{\text{opt}})}. \quad (19)$$

This choice corresponds to the situation when  $\|\hat{u} - u_k\| = 0$ .

**Proposition 1.** *It holds that  $\alpha_{\text{opt}}^{(k+1)} \leq \alpha_{\text{opt}}$ .*

*Proof.* This inequality follows from the monotonicity of  $\psi$ ,  $\lambda$  and the obvious relation

$$\psi\left(\alpha_{\text{opt}}^{(k+1)}\right) \lambda\left(\alpha_{\text{opt}}^{(k+1)}\right) \leq \psi\left(\alpha_{\text{opt}}^{(k+1)}\right) \lambda\left(\alpha_{\text{opt}}^{(k+1)}\right) + \gamma_1\rho\|\hat{u} - u_k\| \lambda\left(\alpha_{\text{opt}}^{(k+1)}\right) = \psi(\alpha_{\text{opt}}) \lambda(\alpha_{\text{opt}}).$$

□

**Lemma 1.** *There is a constant  $c_\psi > 0$  such that*

$$\psi(\alpha_{\text{opt}}) < c_\psi b(\delta).$$

*Proof.* Note that  $b(\delta) = \varphi(\bar{\alpha})$  with  $\bar{\alpha}$  satisfying

$$\sqrt{\bar{\alpha}}\varphi(\bar{\alpha}) = \delta.$$

On the other hand, by definition

$$\sqrt{\alpha_{\text{opt}}}\varphi(\alpha_{\text{opt}}) = \frac{\gamma_2\gamma_\mu}{\gamma_3}\delta.$$

If  $\frac{\gamma_2 \gamma_\mu}{\gamma_3} \leq 1$ , then due to the monotonicity of  $\sqrt{t}\varphi(t)$  we get that  $\alpha_{\text{opt}} \leq \bar{\alpha}$  and

$$\psi(\alpha_{\text{opt}}) = \frac{\gamma_3}{\gamma_\mu} \varphi(\alpha_{\text{opt}}) \leq \frac{\gamma_3}{\gamma_\mu} \varphi(\bar{\alpha}) \leq c_\psi b(\delta),$$

with  $c_\psi = \frac{\gamma_3}{\gamma_\mu}$ . If  $\frac{\gamma_2 \gamma_\mu}{\gamma_3} > 1$  then we have  $\alpha_{\text{opt}} > \bar{\alpha}$  and proceed as

$$\psi(\alpha_{\text{opt}}) = \frac{\gamma_2 \delta}{\sqrt{\alpha_{\text{opt}}}} \leq \frac{\gamma_2 \delta}{\sqrt{\bar{\alpha}}} = c_\psi b(\delta)$$

with  $c_\psi = \gamma_2$ . □

There are two difficulties in choosing  $\alpha_{\text{opt}}^{(k+1)}$ . The first one is connected with the index function  $\psi$  that is usually unknown. The same difficulty appears for linear problems. The second difficulty is due to the nonlinearity. It is connected with a lack of knowledge of the previous iteration error  $\|\hat{u} - u_k\|$ . In recent papers [18, 24] a parameter choice criterion for linear problems has been introduced that adapts automatically to the unknown index function. It appears that this criterion can also be used for adaptation to the unknown term in the left hand side of (18).

As usual, we choose a regularization parameter from the following discrete set

$$\Delta_m^q = \{\alpha_i = \alpha_0 q^i \mid i = 0, 1, \dots, m\}, \quad q > 1. \quad (20)$$

Then, on the  $(k+1)$ -th iteration we compute regularized solutions

$$u_{k+1}^{\alpha_i} := g_{\alpha_i}(A^* A) A^* (y_\delta - G u_k), \quad (21)$$

for each  $\alpha_i \in \Delta_m^q$ . Consider the following two subsets of  $\Delta_m^q$

$$\begin{aligned} M_*^k(\Delta_m^q) &:= \left\{ \alpha \in \Delta_m^q \mid \psi(\alpha) + \gamma_1 \rho \|\hat{u} - u_k\| \leq \frac{\delta}{\lambda(\alpha)} \right\}, \\ M_+^k(\Delta_m^q) &:= \left\{ \alpha_i \in \Delta_m^q \mid \|u_{k+1}^{\alpha_i} - u_{k+1}^{\alpha_j}\| \leq 4 \frac{\delta}{\lambda(\alpha_j)}, j = 0, 1, \dots, i-1 \right\}, \end{aligned} \quad (22)$$

and denote

$$\begin{aligned} \alpha_*^{(k+1)} &:= \max M_*^k(\Delta_m^q), \\ \alpha_+^{(k+1)} &:= \max M_+^k(\Delta_m^q). \end{aligned}$$

The regularization parameter  $\alpha_*^{(k+1)}$  can be considered as the best approximation for  $\alpha_{\text{opt}}^{(k+1)}$  from the set  $\Delta_m^q$ . But in contrast to the parameter  $\alpha_+^{(k+1)}$  it can not be determined. Note, that the error produced by  $u_{k+1}^{\alpha_{\text{opt}}^{(k+1)}}$  is estimated as

$$\|\hat{u} - u_{k+1}^{\alpha_{\text{opt}}^{(k+1)}}\| \leq 2 \left( \psi \left( \alpha_{\text{opt}}^{(k+1)} \right) + \gamma_1 \rho \|\hat{u} - u_k\| \right).$$

We show that the error estimate for  $u_{k+1} := u_{k+1}^{\alpha_+^{(k+1)}}$  is worse only by a constant factor.

**Lemma 2.** *Let  $M_*^k(\Delta_m^q)$  be a proper subset of  $\Delta_m^q$ , i.e.,  $M_*^k(\Delta_m^q) \neq \emptyset$  and  $\Delta_m^q \setminus M_*^k(\Delta_m^q) \neq \emptyset$ . Then it holds that*

$$\|\hat{u} - u_{k+1}\| \leq 3\sqrt{q} \cdot 2 \left( \psi \left( \alpha_{\text{opt}}^{(k+1)} \right) + \gamma_1 \rho \|\hat{u} - u_k\| \right).$$

*Proof.* The proof is based on Theorem 1.1 from [24]. Let  $\{s, t\} \subset \{0, 1, \dots, m\}$  are such indexes that

$$\begin{aligned} \alpha_*^{(k+1)} &= \alpha_s \\ \alpha_+^{(k+1)} &= \alpha_t. \end{aligned}$$

We claim that  $\alpha_s \leq \alpha_t$ . For this, it is sufficient to show that  $M_*^k(\Delta_m^q) \subset M_+^k(\Delta_m^q)$ . Indeed, for any  $\alpha_i > \alpha_j$  such that  $\{\alpha_i, \alpha_j\} \subset M_*^k(\Delta_m^q)$  we have

$$\begin{aligned} \|u_{k+1}^{\alpha_i} - u_{k+1}^{\alpha_j}\| &\leq \|\hat{u} - u_{k+1}^{\alpha_i}\| + \|\hat{u} - u_{k+1}^{\alpha_j}\| \\ &\leq \psi(\alpha_i) + \gamma_1 \rho \|\hat{u} - u_k\| + \psi(\alpha_j) + \gamma_1 \rho \|\hat{u} - u_k\| + \frac{\delta}{\lambda(\alpha_i)} + \frac{\delta}{\lambda(\alpha_j)} \\ &\leq \frac{2\delta}{\lambda(\alpha_i)} + \frac{2\delta}{\lambda(\alpha_j)} \leq \frac{4\delta}{\lambda(\alpha_j)}. \end{aligned}$$

Thus,  $M_*^k(\Delta_m^q) \subset M_+^k(\Delta_m^q)$  and that's why  $\alpha_s \leq \alpha_t$ . From the definition of  $\alpha_s$  it follows that for  $\alpha_{s+1} = q\alpha_s$

$$(\psi(\alpha_{s+1}) + \gamma_1 \rho \|\hat{u} - u_k\|) \lambda(\alpha_{s+1}) > \delta = \left( \psi(\alpha_{\text{opt}}^{(k+1)}) + \gamma_1 \rho \|\hat{u} - u_k\| \right) \lambda(\alpha_{\text{opt}}^{(k+1)}).$$

Since the functions  $\psi$  and  $\lambda$  are increasing, we get that  $\alpha_{s+1} \geq \alpha_{\text{opt}}^{(k+1)}$  and

$$\lambda(\alpha_{\text{opt}}^{(k+1)}) = \frac{1}{\gamma_2} \sqrt{\alpha_{\text{opt}}^{(k+1)}} \leq \frac{1}{\gamma_2} \sqrt{q\alpha_s} = \sqrt{q} \lambda(\alpha_s).$$

Now we get the desired estimate of the iteration error for  $u_{k+1}$

$$\begin{aligned} \|\hat{u} - u_{k+1}\| &= \|\hat{u} - u_{k+1}^{\alpha_t}\| \leq \|\hat{u} - u_{k+1}^{\alpha_s}\| + \|u_{k+1}^{\alpha_s} - u_{k+1}^{\alpha_t}\| \\ &\leq \psi(\alpha_s) + \gamma_1 \rho \|\hat{u} - u_k\| + \frac{\delta}{\lambda(\alpha_s)} + 4 \frac{\delta}{\lambda(\alpha_s)} \\ &\leq 6 \frac{\delta}{\lambda(\alpha_s)} \leq 6\sqrt{q} \frac{\delta}{\lambda(\alpha_{\text{opt}}^{(k+1)})} \\ &= 3\sqrt{q} \cdot 2 \left( \psi(\alpha_{\text{opt}}^{(k+1)}) + \gamma_1 \rho \|\hat{u} - u_k\| \right). \end{aligned}$$

□

**Theorem 1.** Let  $g_\alpha$  fulfill (R1)-(R3). Further, assume that conditions (A1),(A2) are satisfied with  $\rho < \frac{1}{6\gamma_1}$ , where  $\gamma_1$  is the constant from (R1). Consider  $u_{k+1} = g_\alpha(A^*A)A^*(y_\delta - Gu_k)$ ,  $\alpha = \alpha_+^{(k+1)} = \max M_+^k(\Delta_m^q)$  with  $1 < q < \left(\frac{1}{6\gamma_1\rho}\right)^2$ . If  $\hat{u} \in \mathcal{A}_\varphi$  and  $\varphi$  meet (13),(14) then there exist constants  $\sigma \in (0, 1)$ ,  $K \in \mathbb{N}$ ,  $c > 0$  such that:

$$\begin{aligned} \|\hat{u} - u_{k+1}\| &\leq \sigma^{k+1} \|\hat{u} - u_0\|, \text{ for all } k \leq K; \\ \|\hat{u} - u_{k+1}\| &\leq c \cdot b(\delta), \text{ for all } k > K. \end{aligned}$$

*Proof.* Let us introduce

$$B > \frac{6\sqrt{q}\gamma_1\rho}{1-6\sqrt{q}\gamma_1\rho},$$

$$c := 6\sqrt{q}(B+1)c_\psi,$$

$$\sigma := 6\sqrt{q} \frac{B+1}{B} \gamma_1 \rho < 1,$$

where  $c_\psi$  is the constant from Lemma 1. We show that for such  $c$  and  $\sigma$  the statement of the theorem holds true.

As long as the iteration error  $\|\hat{u} - u_k\|$  of  $k$ -th iteration is so large that it holds

$$\psi(\alpha_{\text{opt}}^{(k+1)}) \leq \frac{1}{B} \gamma_1 \rho \|\hat{u} - u_k\|,$$

the iteration error of the next iteration can be estimated as

$$\begin{aligned} \|\hat{u} - u_{k+1}\| &\leq 6\sqrt{q} \left( \psi(\alpha_{\text{opt}}^{(k+1)}) + \gamma_1 \rho \|\hat{u} - u_k\| \right) \\ &\leq 6\sqrt{q} \frac{B+1}{B} \gamma_1 \rho \|\hat{u} - u_k\| \\ &= \sigma \|\hat{u} - u_k\|, \end{aligned}$$

and we observe error decrease.

Let  $K$  be the iteration number when for the first time

$$\psi(\alpha_{\text{opt}}^{(K)}) > \frac{1}{B} \gamma_1 \rho \|\hat{u} - u_{K-1}\|.$$

Then in view of Proposition 1 and Lemma 1 the iteration error  $\|\hat{u} - u_K\|$  is bounded by

$$\begin{aligned} \|\hat{u} - u_K\| &\leq 6\sqrt{q} \left( \psi(\alpha_{\text{opt}}^{(K)}) + \gamma_1 \rho \|\hat{u} - u_{K-1}\| \right) \\ &\leq 6\sqrt{q}(B+1) \psi(\alpha_{\text{opt}}^{(K)}) \leq c \cdot b(\delta). \end{aligned} \tag{23}$$

Now we show by induction that for all  $l \geq K$  the bound (23) holds. The basis of induction for  $l = K$  holds true. For some  $l > K$  assume

$$\|\hat{u} - u_l\| \leq c \cdot b(\delta).$$

Then for the error of the  $(l + 1)$ -th iteration we have the estimate

$$\begin{aligned} \|\hat{u} - u_{l+1}\| &\leq 6\sqrt{q} \left( \psi \left( \alpha_{\text{opt}}^{(l+1)} \right) + \gamma_1 \rho \|\hat{u} - u_l\| \right) \\ &\leq 6\sqrt{q} ( \psi(\alpha_{\text{opt}}) + B \sigma c_\psi b(\delta) ) \\ &\leq c \cdot b(\delta), \end{aligned}$$

as claimed.  $\square$

## 4 A-priori Choice of the Regularization Parameter

Theorem 1 guarantees a stable behavior of the iteration process (9) under two conditions, where the first one is connected with the constant  $\rho$  in (A2). For Tikhonov-Phillips regularization, when  $\gamma_1 = 1$ , this condition reads as  $0 < \rho < 1/6$ . It is interesting to note that if  $A$  has a bounded inverse then the convergence of (8) can be guaranteed under the mild assumption that  $0 < \rho < 1$ . In this case

$$\|\hat{u} - u_{k+1}\| \leq \delta \|A^{-1}\| \frac{1 - \rho^{k+1}}{1 - \rho} + \rho^{k+1} \|\hat{u} - u_0\|.$$

Thus, the stronger requirement  $\rho < 1/6$  could be seen as a “payment” for dealing with an ill-posed problem. But as we will see below, it is just a payment for the lack of knowledge of the solution smoothness, or a payment for adaptation.

Indeed, if the index function  $\varphi$  in (12) is known, then one can take *a-priori* any  $\alpha = \tilde{\alpha}$  such that

$$\beta_1 \frac{\delta}{\lambda(\tilde{\alpha})} \leq \psi(\tilde{\alpha}) \leq \beta_2 \psi(\alpha_{\text{opt}}), \quad (24)$$

where  $\beta_1, \beta_2 > 0$  do not depend on  $\delta$ . If this value of the regularization parameter is fixed for all iterations, then the iteration process (9) exhibits the same behavior as in Theorem 1 under the assumption that  $0 < \rho < \frac{1}{\gamma_1}$ . For Tikhonov-Phillips regularization it reads as  $0 < \rho < 1$ .

**Theorem 2.** Assume  $g_\alpha$  fulfill (R1)-(R3). Further, assume that conditions (A1),(A2) are satisfied with  $\rho < \frac{1}{\gamma_1}$ , where  $\gamma_1$  is the constant from (R1). Consider  $u_{k+1} = g_{\tilde{\alpha}}(A^*A)A^*(y_\delta - Gu_k)$ , where  $\tilde{\alpha}$  satisfies (24). If  $\hat{u} \in \mathcal{A}_\varphi$  and  $\varphi$  meet (13),(14), then there exist constants  $\sigma \in (0, 1)$ ,  $K \in \mathbb{N}$ ,  $c > 0$  such that:

$$\begin{aligned} \|\hat{u} - u_{k+1}\| &\leq \sigma^{k+1} \|\hat{u} - u_0\|, \text{ for all } k \leq K; \\ \|\hat{u} - u_{k+1}\| &\leq c \cdot b(\delta), \text{ for all } k > K. \end{aligned}$$

*Proof.* Let us introduce

$$\begin{aligned} B &> \frac{\gamma_1 \rho}{1 - \gamma_1 \rho}, \\ c &:= (B + 1) \left( 1 + \frac{1}{\beta_1} \right) \beta_2 c_\psi, \\ \sigma &:= \frac{B+1}{B} \gamma_1 \rho < 1, \end{aligned}$$

where  $\beta_1, \beta_2$  are from (24) and  $c_\psi$  is from Lemma 1. We show that with such chosen constants the statement of the theorem follows. Indeed, as long as

$$\psi(\tilde{\alpha}) + \frac{\delta}{\lambda(\tilde{\alpha})} \leq \frac{1}{B} \gamma_1 \rho \|\hat{u} - u_k\|,$$

we have the estimate

$$\begin{aligned} \|\hat{u} - u_{k+1}\| &\leq \psi(\tilde{\alpha}) + \frac{\delta}{\lambda(\tilde{\alpha})} + \gamma_1 \rho \|\hat{u} - u_k\| \\ &\leq \sigma \|\hat{u} - u_k\|, \end{aligned}$$



and error decrease follows. Now let  $K$  be the iteration number when for the first time

$$\psi(\tilde{\alpha}) + \frac{\delta}{\lambda(\tilde{\alpha})} > \frac{1}{B} \gamma_1 \rho \|\hat{u} - u_k\|.$$

Then we have the estimate

$$\begin{aligned} \|\hat{u} - u_K\| &\leq \psi(\tilde{\alpha}) + \frac{\delta}{\lambda(\tilde{\alpha})} + \gamma_1 \rho \|\hat{u} - u_{K-1}\| \\ &\leq c \cdot b(\delta). \end{aligned}$$

Similarly to the proof of Theorem 1 we obtain the desired bound for all  $k > K$  by induction.  $\square$

*Remark 1.* We would like to note an important consequence of the above result for severely ill-posed problems. This is of special interest for us, since we will consider an application of the iterative process (9) to one of the inverse heat conduction problems, which are known to be severely ill-posed. For such problems  $\psi(\alpha) \asymp \log^{-\mu} \frac{1}{\alpha}$  with some  $\mu > 0$ , which may be unknown. As it follows from [11], even with unknown  $\mu$ , the regularization parameter chosen a-priori as

$$\alpha = c_2 \delta^p, \quad c_2 > 0, \quad 0 < p < 2,$$

meets (24) and gives the order optimal accuracy.

## 5 Initial Temperature Reconstruction for a Nonlinear Heat Equation

We now give an example of a nonlinear inverse problem where a splitting of the type (7) is quite natural. Consider the heat equation

$$\frac{\partial T}{\partial t} = \frac{\partial^2 T}{\partial z^2} - B(T), \quad (25)$$

which describes the evolution of the temperature  $T(z, t)$  in some material in one space dimension  $z \in [0, 1]$  during the time  $t \in (0, t_f]$ . The nonlinear function  $B(T)$  represents the heat source or sink. Equation (25) is additionally equipped with boundary and initial conditions

$$\begin{aligned} \frac{\partial T}{\partial z} &= 0 \quad z \in \{0, 1\}, \quad t \in (0, t_f], \\ T(z, 0) &= u(z) \quad z \in [0, 1]. \end{aligned} \quad (26)$$

In many cooling processes of hot materials, e.g., in glass cooling [25], polymer processing [21, 22], the initial temperature is inaccessible. The only available data is the boundary temperature  $y$  at some part of the boundary, i.e.,

$$y(t) = T(0, t), \quad t \in [0, t_f].$$

Thus, letting  $F$  be the operator that maps the initial temperature  $u$  into the boundary temperature  $y$ , we obtain an inverse problem of the type (1). We would like to note that this inverse problem of the initial temperature reconstruction *only* from boundary measurements has been relatively ignored in the literature. In the case of the linear heat equation, we are aware only of the engineering works [21, 22]. As it is pointed out in [21], in the existing publications concerning the initial temperature reconstruction one usually uses some additional temperature measurements inside the material. However, in the above mentioned applications such data is unavailable. To our knowledge, [25] is the first work where this inverse problem is considered for a nonlinear heat equation. The approach proposed in [25] consists in applying a splitting of the operator  $F$  of the type (7). To present this splitting, consider for a moment a linear heat equation

$$\frac{\partial T}{\partial t} = \frac{\partial^2 T}{\partial z^2} - b(z, t) \quad (27)$$

with conditions (26). Let us examine an operator  $L$  that maps both  $u$  and  $b$  into the boundary temperature  $y$ . Then, the inverse problem for  $u$  reads as

$$L(u, b) = y. \quad (28)$$

This problem is still nonlinear, but using the affine decomposition [6]

$$L(u, b) = L(u, 0) + L(0, b), \quad (29)$$

one obtains the linear problem for  $u$

$$L(u, 0) = y - L(0, b).$$

Note, that the nonlinear source term  $B$  in (25) can be considered as dependent on the initial temperature. Indeed, let  $S$  denote the solution operator of equation (25) that maps  $u$  into  $T$ . Then,

$$B(T) = B(S(u)) =: \mathbf{B}(u).$$

Thus, the affine decomposition (29) suggests the following splitting of  $F$

$$\begin{aligned} Fu = L(u, \mathbf{B}(u)) &= L(u, 0) + L(0, \mathbf{B}(u)) \\ &= Au + Gu, \end{aligned}$$

with

$$\begin{aligned} Au &:= L(u, 0), \\ Gu &:= L(0, \mathbf{B}(u)). \end{aligned} \quad (30)$$

For checking the numerical performance of the iterative process (9), we take a specific example of the heat equation (25) that models cooling of a hot glass plate. It is well-known that for processes involving high temperatures, heat radiation needs to be taken into account. We consider the situation when the glass can be assumed to be a grey material. Then, equation (25) has the following form (see, e.g., [19, p. 714])

$$\frac{\partial T}{\partial t} = \frac{\partial^2 T}{\partial z^2} - \kappa \left( 4\sigma T^4 - 2\pi \int_{[-1,1]} Id\mu \right), \quad (31)$$

where  $\kappa$  is the absorption coefficient — material characteristic responsible for the influence of radiation effects;  $\sigma = 5.67051 \cdot 10^{-8}$  is the dimensionless Stefan-Boltzmann constant. The function  $I = I(z, \mu)$  is the so-called radiative intensity that depends on the position  $z$  and on the direction coordinate along the plate thickness  $\mu \in [-1, 1]$ . This function is related to the temperature through the radiative transfer equation that for glass cooling processes [15] can be taken as

$$\mu \frac{\partial I}{\partial z} = \kappa \left( \frac{\sigma}{\pi} T^4 - I \right). \quad (32)$$

We assume that it is equipped with the so-called black surface boundary conditions [19, p. 309]

$$\begin{aligned} I(0, \mu) &= \frac{\sigma}{\pi} T_a^4, \quad \mu > 0 \\ I(1, \mu) &= \frac{\sigma}{\pi} T_a^4, \quad \mu < 0 \end{aligned} \quad (33)$$

where  $T_a$  is the ambient temperature.

Although equations (31) and (32) are coupled, equation (31) can still be seen as a particular case of (25). The source term  $B$  in this case depends on the temperature in a nonlinear and especially nonlocal way. In the Appendix we outline a numerical treatment of (31)-(32) that is used for calculating  $Fu$  and  $Gu$ .

It is well-known, that inverse problems associated with parabolic problems are severely (exponentially) ill-posed [6, 11]. Thus, in our case each iteration (8) is a severely ill-posed operator equation. According to Example 1, Tikhonov-Phillips regularization is appropriate for such problems. That's why in the regularized procedure (9) we take  $g_\alpha(t) = \frac{1}{\alpha+t}$ . Then (9) can be alternatively written as

$$(\alpha I + A^* A) u_{k+1} = A^* (y_\delta - Gu_k). \quad (34)$$

The numerical solution of equation (34) requires its discretization. We will use a Galerkin method for this purpose. It requires a collection of finite dimensional subspaces  $\{X_N\}$  such that  $\dim X_N = N$  and  $\bigcup_{N=1}^{\infty} X_N = X$ . For the problem under consideration, the Hilbert space  $X$

consists of functions  $u$  that are used as initial conditions in (26). We take  $X = L^2(0, 1)$ . Let  $\{\varphi_i\}_{i=1}^N$  be a basis of  $X_N$ . For a fixed discretization parameter  $N$ , the Galerkin approximate solution  $u_{k+1}^N = \sum_{j=1}^N c_j \varphi_j$  of (34) is found by solving the linear system

$$\alpha \sum_{j=1}^N c_j (\varphi_j, \varphi_i) + \sum_{j=1}^N c_j (A\varphi_j, A\varphi_i) = (y_\delta - G u_k, A\varphi_i), \quad i = 1, \dots, N. \quad (35)$$

In our numerical experiments we increase  $N$  until some convergence criterion is satisfied, e.g.,

$$\frac{\|u_{k+1}^N - u_{k+1}^{N-1}\|}{\|u_{k+1}^{N-1}\|} < \text{tol}, \quad (36)$$

where  $\text{tol} > 0$  is a chosen tolerance.

The choice of the basis  $\{\varphi_i\}$  is suggested by the special structure of the operator  $A$ , namely its relation to (27). It is easy to see that for  $u(z) = \cos(\pi(i-1)z)$ ,  $b(z, t) \equiv 0$  the solution of (27) is  $T(z, t) = \cos(\pi(i-1)z) \cdot e^{-\pi^2(i-1)^2 t}$ . Thus,

$$A[\cos(\pi(i-1)z)] = e^{-\pi^2(i-1)^2 t}.$$

Clearly,  $\varphi_i(z) = \cos(\pi(i-1)z)$  is a basis in  $L^2(0, 1)$ .

## 5.1 Numerical results

As it was pointed out in [17] for discretized problems one should consider the following modification of the set  $M_+^k(\Delta_m^q)$  from (22)

$$\tilde{M}_+^k(\Delta_m^q) := \left\{ \alpha_i \in \Delta_m^q \mid \|u_{k+1}^{\alpha_i} - u_{k+1}^{\alpha_j}\| \leq 4\varepsilon \frac{\delta}{\lambda(\alpha_j)}, \quad j = 0, 1, \dots, i-1 \right\},$$

where  $\varepsilon$  is a *tuning parameter* that depends on the problem and needs to be chosen by the user. Making a lot of numerical experiments with different initial conditions and noise levels, we discovered that for our problem  $\varepsilon = 0.07$  is a good choice.

Now we present the numerical results for one initial temperature profile. Such a profile arises frequently in practice when a piece of glass is cooled uniformly on both boundaries. It is constructed in two steps. First, we consider the function

$$\hat{u}^{\text{scl}}(z) = 1.5 + \sum_{i=1}^8 -\frac{8}{(2i)^4} \cos(\pi(2i)z).$$

Then we scale it such that its range is in the interval  $[800, 1200]$ , i.e.,

$$\hat{u}^{\text{scl}} = 800 + (\hat{u}^{\text{scl}} - \min(\hat{u}^{\text{scl}})) \frac{400}{\max(\hat{u}^{\text{scl}}) - \min(\hat{u}^{\text{scl}})}.$$

The considered range is a typical temperature range in Kelvin that appears in glass production.

Numerical results are presented in Figure 1 for two different noise levels 1% and 0.1% (description of the noise generation procedure is presented in Appendix). The monotone decrease of the iteration error corresponds to Theorem 1. It was also observed, that the error remains the same if the iterative process is performed further, which is predicted as well. We would like to note the small number of iterations needed to achieve the stable stage of the process. The final approximation is improved with the decrease of the noise.

## 5.2 Validation of assumptions by a computational experiment

One observes that there is quite strong and complicated coupling between (31) and (32). This makes an analysis of the corresponding nonlinear operators  $F$  and  $G$  extremely complicated. That's why there is little hope that assumptions (A1)-(A2) can be checked analytically, and especially to obtain a sharp estimate of the constant  $\rho$  in (A2). This situation is quite frequent in inverse problems arising from real applications (see for example remark in [8, p. 284]). In a view of this, we decided

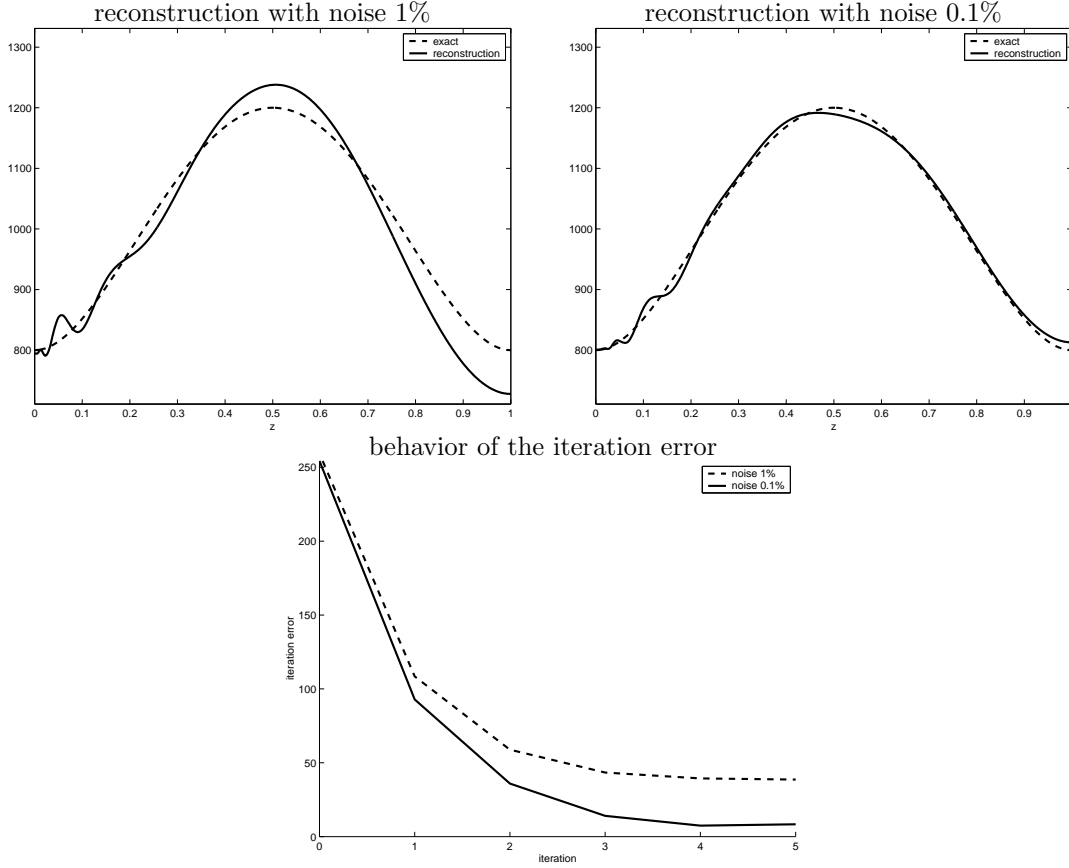


Figure 1: Reconstruction of  $\hat{u}$  for different noise levels.

to try the approach of the so-called “*computational experiment*” introduced in the late 80-th by the Russian mathematician K. I. Babenko (see, e.g., [1, 2]). The idea is to design a numerical procedure which could show that assumptions (A1)-(A2) are reasonable as well as giving an estimate on the constant  $\rho$ . We propose the following procedure:

1. Generate randomly a certain amount  $M$  of function pairs  $\{(u_1^{(j)}, u_2^{(j)}), j = 1 \dots M\}$ .
2. For each pair find  $v^{(j)}$  such that

$$Av^{(j)} = Gu_1^{(j)} - Gu_2^{(j)}. \quad (37)$$

3. Compute

$$\rho_j = \frac{\|v^{(j)}\|}{\|u_1^{(j)} - u_2^{(j)}\|}.$$

4. Finally, the quantity  $\tilde{\rho} = \max_j \rho_j$  can be considered as an estimate of  $\rho$  in (A2).

In this procedure the most difficult part is step 2. In our case, the inverse of the operator  $A$  is severely ill-posed. This makes a numerical solution of  $Au = y$ , even with unnoisy data  $y$ , impossible without regularization due to round-off errors. In the case of (37), additional errors will appear due to the discretization of  $G$  and these errors will be unknown. As it was pointed out in [24], the noise-free analog of the adaptive criterion used in Section 3 corresponds to the so-called quasi-optimality criterion, which was first introduced in [32]. It consists in choosing a regularization parameter  $\alpha_{\text{qo}}^{(k+1)} := \alpha_j$  such that

$$\|u_{k+1}^{\alpha_j} - u_{k+1}^{\alpha_{j-1}}\| = \min\{\|u_{k+1}^{\alpha_i} - u_{k+1}^{\alpha_{i-1}}\|, i = 1, \dots, m\}. \quad (38)$$

We will use this criterion for choosing the regularization parameter in (37).

Let us describe now a possible realization of the step 1. The random functions will be taken from the set of piecewise linear splines with  $N_d$  equidistant knots  $\mathcal{K} := \left\{ \frac{i}{N_d-1}, i = 0, \dots, N_d - 1 \right\}$ . Since in the noise-free case all iterative solutions have the same value at  $z = 0$  ( $u_k(0) = y(0)$ ), the assumption (A1) is sufficient to check for functions  $\hat{u}(0) = u(0)$ . That's why the spline set will be also considered with this property

$$S(N_d, T_{\text{mid}}, T_{\text{step}}, N_{\text{step}}) = \left\{ s \text{ is a piecewise linear spline with knots } \mathcal{K} \mid \right. \\ \left. s(0) = T_{\text{mid}}, s\left(\frac{i}{N_d-1}\right) \in \{T_{\text{mid}} + j \cdot T_{\text{step}}, j = -N_{\text{step}}, \dots, N_{\text{step}}\}, \right. \\ \left. i = 1, \dots, N_d - 1 \right\}.$$

Then random choice of a function  $s_{\text{rnd}}$  from this set is realized as

$$s_{\text{rnd}}\left(\frac{i}{N_d-1}\right) = T_{\text{mid}} + \xi_i T_{\text{step}}, \quad i = 1, \dots, N_d - 1,$$

where  $\{\xi_i\}$  are independent random variables uniformly distributed over  $\{-N_{\text{step}}, \dots, N_{\text{step}}\}$ .

In our computations we consider the following values of the involved parameters:  $M = 800$ ,  $N_d = 5$ ,  $T_{\text{mid}} = 1000$ ,  $T_{\text{step}} = 100$ ,  $N_{\text{step}} = 3$ . Let us also note, that the parameter  $\kappa$  in (31) has a crucial influence on the operators  $F$  and  $G$  — the bigger it is, the bigger is the difference between operator  $F$  and its linearization  $A$ . We did computations for three different values of  $\kappa$ , namely — 0.1, 0.5, 1, and observe how the estimate of  $\rho$  depends on them. The discretization of the operator  $G$  was chosen fine enough to obtain numerical results in some reasonable time (see Section A.3).

One random pair  $(u_1, u_2)$ , as well as corresponding  $G u_1 - G u_2$ ,  $v$  and  $A v$  are presented in Figure 2. The numerical values obtained in this random realization are gathered in Table 1. Note the good correspondence between  $A v$  and  $G u_1 - G u_2$  in Figure 2(d). All computed  $\{\rho_j\}$  for different  $\kappa$  are presented in Figure 3. One observes that for  $\kappa = 0.1$  the estimate  $\hat{\rho}$  is less than  $1/6$  that suggests the applicability of Theorem 1, whereas for other values of  $\kappa$  Theorem 2 with Remark 1 can be used.

## 6 Final Remarks

First of all, we note that satisfactory numerical results were obtained in cases when the constant  $\rho$  in (A2) is bigger then required by Theorem 1. In particular, the results in Section 5.1 were obtained in the case when the estimate of  $\rho$  according to the computational experiment is bigger then  $1/6$ . This is not a contradiction, because Theorem 1 guarantees the order-optimal convergence rate for small enough  $\rho$  and for bigger  $\rho$  the iteration error can still have a stable behavior but possibly not the order-optimal convergence. In the future work we will study this situation in more detail. Note that for severely ill-posed problems the theoretical justification of the regularization procedure (34) with fixed  $\alpha$  is given by Theorem 2 and Remark 1 under the much milder assumption  $\rho < 1$ .

Another point is connected with further applications of the proposed method. Similar splittings can be done for the reconstruction of the initial temperature in the nonlinear backward heat equation. We would like to stress once again the derivative-free nature of the proposed iterative method. In the considered example the use of the Fréchet derivative in computations is rather complicated due to the implicit definition of the forward operator. But even for problems where the Fréchet derivative is known, its iterative update and regularized inversion is very time consuming. Whereas in the proposed method, one needs to invert fixed linear operators only once. The advantage of the much easier implementation is clear in this case. However, it is not clear whether the overall time needed to achieve a desired approximation by Newton-like methods is bigger than by the proposed one, because the above mentioned inversion of the linear operators should be done for several values of the regularization parameter. Hence, it is interesting to apply our method for nonlinear inverse problems with known Fréchet derivative, and to compare it with derivative-based methods.

The implementation of the considered technique for multidimensional problems is straightforward. However, in this case a number of difficulties arise. In general, it can not be expected that the functions  $\{A \varphi_i\}$  in (35) will be known exactly. As it is pointed out in [25], one should be concerned with the choice of the discretization parameter in this case. Also, in multidimensional problems one needs to deal with a larger amount of basis functions that increases the size of matrices in

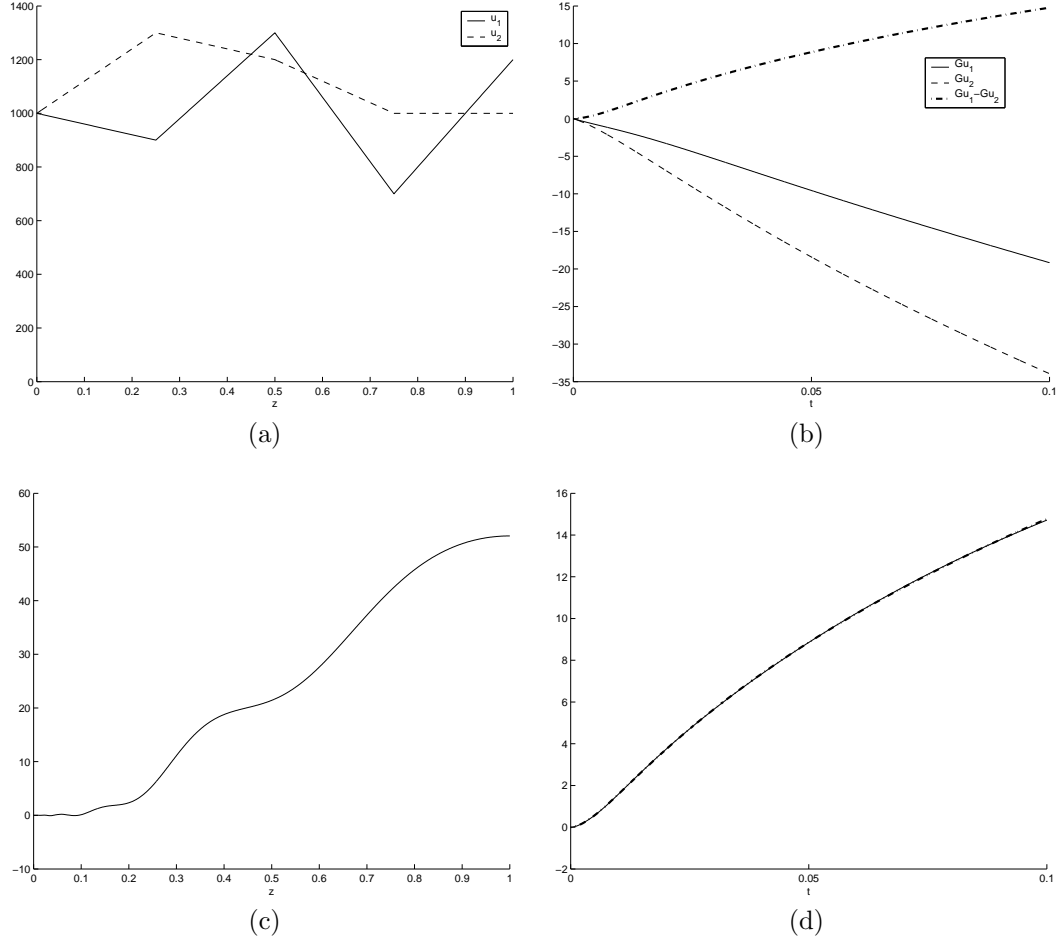


Figure 2: Example of one random realization in the computational experiment: (a) – random functions  $u_1$  and  $u_2$ ; (b) – corresponding images  $Gu_1$  and  $Gu_2$ ; (c) – approximate solution  $v$  of the equation (37); (d) – functions  $Av$  (solid line) and  $Gu_1 - Gu_2$  (dash-dot line). Calculations were done with  $\kappa = 0.1$

$$\begin{aligned}\|u_1 - u_2\| &= 189.2969 \\ \|Gu_1 - Gu_2\| &= 2.9564 \\ \|v\| &= 30.1724 \\ \rho &= 0.15939\end{aligned}$$

Table 1: Numerical values obtained in the random realization from Figure 2.

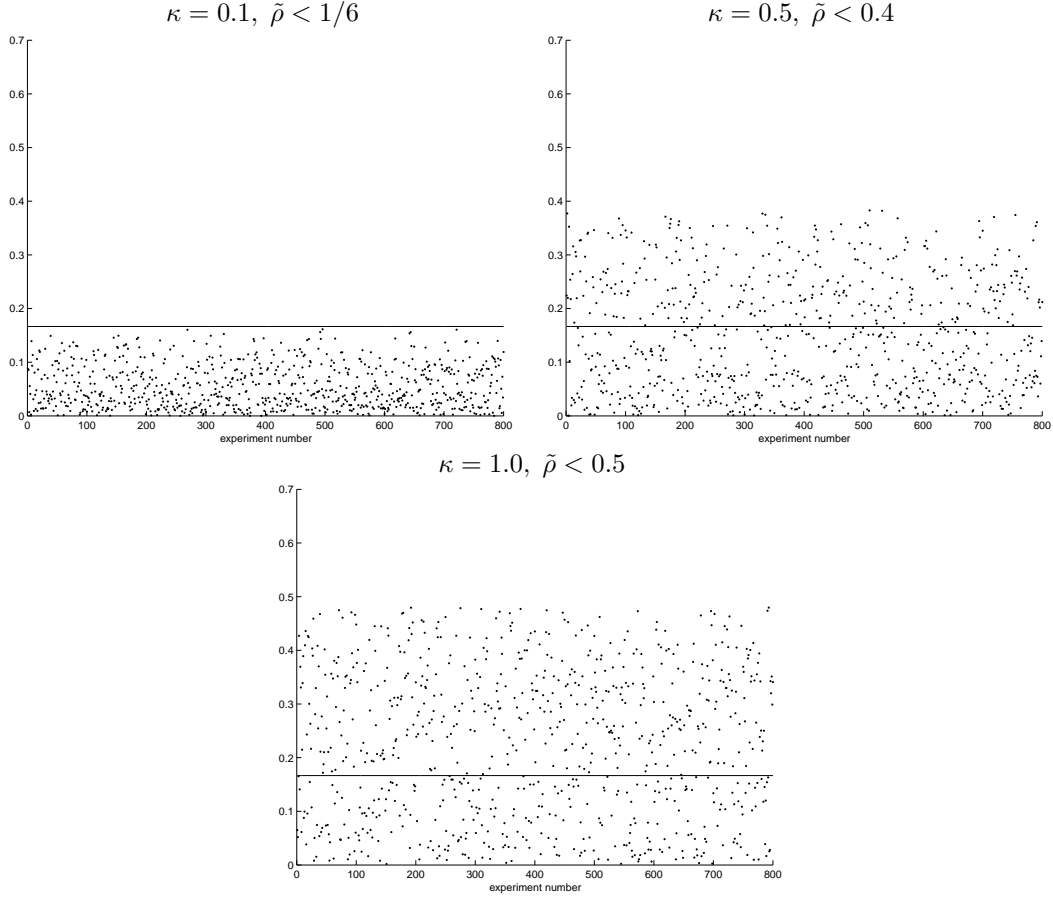


Figure 3: Computed  $\{\rho_j\}$  for different  $\kappa$ . Line corresponds to the value  $1/6$ .

(35). This complicates the application of direct methods to (35). Therefore, in multidimensional problems one probably should apply iterative methods for computing discretized solutions.

## 7 Acknowledgments

The authors acknowledge support from the DFG via contract PI 408/3-1. Further, the authors would like to thank Prof. Heinz W. Engl and Dr. Martin Burger for their helpful comments concerning this work during a visit of the Radon Institute for Computational and Applied Mathematics (Linz, Austria).

## A Appendix

Let us briefly outline the numerical procedure that was used to solve equations (31)-(32). First, we show how the source term in (31) can be computed for a given temperature  $T$ . For this purpose, we consider the approximation of the integral  $\int_{[-1,1]} \cdot d\mu$  by the so-called Discrete Ordinate Method ( $S_p$ -approximation) [19, Ch. 15], which consists in choosing  $\{\mu_i, \omega_i\}_{i=1}^p$  such that

$$\int_{[-1,1]} I(z, \mu) d\mu \approx \sum_{i=1}^p I(z, \mu_i) \omega_i.$$

The functions  $I(z, \mu_i)$  are determined from the equations

$$\begin{aligned} \mu_i \frac{\partial I(z, \mu_i)}{\partial z} &= \kappa \left( \frac{\sigma}{\pi} T^4 - I(z, \mu_i) \right), \quad i = 1, \dots, p \\ I(0, \mu_i) &= \frac{\sigma}{\pi} T_a^4, \quad \mu_i > 0, \\ I(1, \mu_i) &= \frac{\sigma}{\pi} T_a^4, \quad \mu_i < 0. \end{aligned} \tag{39}$$

Thus, each equation (39) is a linear hyperbolic partial differential equation. Its numerical solution can be done by a finite difference approximation of the derivative  $\frac{\partial}{\partial z}$  for a given  $T$ .

Having a numerical procedure to compute the source term in (31) for a given temperature  $T$ , the whole solution of (31) can be made by the standard finite difference method with explicitly incorporated nonlinear source term. In this case, one deals with space and time discretizations

$$\begin{aligned} \left\{ z_i = \frac{i}{N_z}, \quad i = 0, \dots, N_z \right\} \\ \left\{ t_j = \frac{j}{N_t} t_f, \quad j = 0, \dots, N_t \right\} \end{aligned}$$

and obtains a discretized solution  $\tilde{T}(z_i, t_j)$ . Consequently, the discretized approximation  $\tilde{y}$  of the  $y = Fu$  is a piece-wise linear spline with knots  $\{t_j\}$  and values

$$\tilde{y}(t_j) = \tilde{T}(0, t_j).$$

The approximation of  $Gu$  was done by first running a numerical procedure for (31) with  $u$  in the initial condition. This gives the discretized source term. Then we run a numerical procedure for (27) with that discretized source term and zero as initial condition.

### A.1 Generation of the noise

For a given initial temperature  $\hat{u}$  we obtain a discretized approximation  $\tilde{y}$  of  $F\hat{u}$  that is known in  $(N_t + 1)$  points. Corresponding noisy data  $y_\delta$  is obtained from  $\tilde{y}$  by the following steps:

- first, consider  $\tilde{y}$  on a more coarse time discretization  $N_t^c = N_t/k^c$ , which gives  $\tilde{y}^c$  defined as

$$\tilde{y}^c(t_{j \, k^c}) = \tilde{y}(t_{j \, k^c}), \quad j = 0, \dots, N_t^c;$$

- then perturb  $\tilde{y}^c$  with random noise

$$y_\delta(t_{j \, k^c}) = \tilde{y}^c(t_{j \, k^c}) + \delta_r \cdot \xi_j,$$

where  $\delta_r > 0$  is a noise level and  $\{\xi_j\}$  are independent random variables with uniform distribution over  $[-1, 1]$ ;

- the final noisy data  $y_\delta$  is considered to be a piece-wise linear spline with knots  $\{t_{j \, k^c}, \quad j = 0, \dots, N_t^c\}$ .



## A.2 Parameters used in the simulations

For the numerical results presented in Section 5.1 the following model parameters were used:  $t_f = 0.1$ ,  $\kappa = 0.5$ ,  $T_a = 300$  (temperature in Kelvin). For the approximation of the nonlinear operators  $F$  and  $G$ , space and time discretization were taken as  $N_z = 400$ ,  $N_t = 400$ ; in the Discrete Ordinate Method we used  $p = 8$  (see [19, p. 546] for corresponding values of  $\{\mu_i, \omega_i\}$ ).

For the discrete set of regularization parameters (20) we used  $\alpha_0 = 10^{-10}$ ,  $q = 2$ ,  $m = 25$ . The value for the tolerance in (36) was  $\text{tol} = 10^{-5}$ .

For the noise generation we used  $k_c = 4$  and two noise levels  $\delta_r \in \{10, 1\}$ . Since the average temperature of the used function  $\hat{u}$  is 1000, these noise levels can be considered as 1% and 0.1% respectively.

## A.3 Parameters used in the computational experiment

We used  $N_z = 201$ ,  $N_t = 1601$  as discretization parameters for the results in Section 5.2. All other parameters, except of  $\kappa$ , were taken as in previous subsection.

## References

- [1] K. I. Babenko, V. Y. Petrovich, and A. I. Rakhmanov. A computational experiment in the theory of surface wave of finite amplitude. *Sov. Math. Dokl.*, 38:327–331, 1989.
- [2] K. I. Babenko and M. M. Vasil’ev. Demonstrative calculations in the theory of hydrodynamic stability. *Sel. Math. Sov.*, 9:171–197, 1990.
- [3] A. B. Bakushinskii. Remarks on choosing regularization parameter using the quasi-optimality and ratio criterion. *USSR Comp. Math. Math. Phys.*, 24:181–182, 1984.
- [4] A. B. Bakushinskij. The problem of the convergence of the iteratively regularized Gauss-Newton method. *Comput. Math. Math. Phys.*, 32:1353–1359, 1992.
- [5] B. Blaschke, A. Neubauer, and O. Scherzer. On convergence rates for the iteratively regularized Gauss-Newton method. *IMA J. Numer. Anal.*, 17:421–436, 1997.
- [6] S. Dahlke and P. Maaß. An outline of adaptive wavelet Galerkin methods for Tikhonov regularization of inverse parabolic problems. In Y.-C. Hon and et al, editors, *Recent Development in Theories and Numerics. Proceedings of the International Conference on Inverse Problems*, pages 56–66, Hong Kong, China, 9-12 January 2002.
- [7] P. Deuffhard, H. W. Engl, and O. Scherzer. A convergence analysis of iterative methods for the solution of nonlinear ill-posed problems under affinity invariant conditions. *Inverse Problems*, 14:1081–1106, 1998.
- [8] H. W. Engl, M. Hanke, and A. Neubauer. *Regularization of inverse problems*. Kluwer Academic Publishers, Dordrecht, 1996.
- [9] H. W. Engl, K. Kunisch, and A. Neubauer. Convergence rates for Tikhonov regularisation of nonlinear ill-posed problems. *Inverse Problems*, 5:523–540, 1989.
- [10] B. Hofmann. *Mathematik inverser Probleme*. Teubner, Stuttgart, 1999.
- [11] B. Hofmann and M. Yamamoto. Convergence rates for Tikhonov regularization based on range inclusions. *Inverse Problems*, 21:805–820, 2005.
- [12] T. Hohage. Logarithmic convergence rates of the iteratively regularized Gauss-Newton method for an inverse potential and an inverse scattering problem. *Inverse Problems*, 13:1279–1299, 1997.
- [13] B. Kaltenbacher. Some Newton-type methods for the regularization of nonlinear ill-posed problems. *Inverse Problems*, 13:729–753, 1997.
- [14] A. Kirsch. *An introduction to the mathematical theory of inverse problems*. Springer-Verlag, Berlin, 1996.

- [15] F. T. Lentes and N. Siedow. Three-dimensional radiative heat transfer in glass cooling processes. *Glass Science and Technology – Glastechnische Berichte*, 72:188–196, 1999.
- [16] B. A. Mair. Tikhonov regularization for finitely and infinitely smoothing operators. *SIAM J. Math. Anal.*, 25:135–147, 1994.
- [17] P. Mathé and S. V. Pereverzev. Discretization strategy for linear ill-posed problems in variable hilbert scales. *Inverse Problems*, 19:1263–1277, 2003.
- [18] P. Mathe and S. V. Pereverzev. Geometry of linear ill-posed problems in variable Hilbert scales. *Inverse Problems*, 19:789–803, 2003.
- [19] M. F. Modest. *Radiative Heat Transfer*. McGraw-Hill Inc., 1993.
- [20] A. Neubauer. Tikhonov regularisation for nonlinear ill-posed problems: Optimal convergence rates and finite-dimensional approximation. *Inverse Problems*, 5:541–557, 1989.
- [21] K. T. Nguyen and A. Bendata. An inverse approach for the prediction of the temperature evolution during induction heating of a semi-solid casting billet. *Modelling Simul. Mater. Sci. Eng.*, 8:857–870, 2000.
- [22] K. T. Nguyen and M. Prystay. An inverse method for estimation of the initial temperature profile and its evolution in polymer processing. *Int. J. Heat Mass Transfer*, 42:1969–1978, 1999.
- [23] S. Pereverzev and E. Schock. Morozov’s discrepancy principle for Tikhonov regularization of severely ill-posed problems in finite-dimensional subspaces. *Numer. Funct. Anal. Optim.*, 21:901–916, 2000.
- [24] S. Pereverzev and E. Schock. On the adaptive selection of the parameter in regularization of ill-posed problems. *Schriften zur Funktionalanalysis und Geomathematik 1*, Department of Mathematics, TU Kaiserslautern, 2003. to appear in *SIAM J. Numer. Anal.*
- [25] S. S. Pereverzyev, R. Pinnau, and N. Siedow. Initial temperature reconstruction for a nonlinear heat equation: application to radiative and conductive heat transfer. In D. Lesnic, editor, *5th International Conference on Inverse Problems in Engineering: Theory and Practice*, volume III, pages P02: 1–8, Leeds, UK, 2005. Leeds University Press.
- [26] A. Rieder. *Keine Probleme mit inversen Problemen. Eine Einführung in ihre stabile Lösung*. Vieweg, Wiesbaden, 2003.
- [27] E. Schock. Approximate solution of ill-posed equations: Arbitrarily slow convergence vs. superconvergence. In G. Hämmerlein and K. H. Hoffmann, editors, *Constructive methods for the practical treatment of integral equations*, pages 234–243, Basel, 1985. Birkhäuser.
- [28] T. Schröter and U. Tautenhahn. Error estimates for Tikhonov regularization in Hilbert scales. *Numer. Funct. Anal. Optim.*, 15:155–168, 1994.
- [29] U. Tautenhahn. On a general regularization scheme for nonlinear ill-posed problems. *Inverse Problems*, 13:1427–1437, 1997.
- [30] U. Tautenhahn. Optimality for ill-posed problems under general source conditions. *Numer. Funct. Anal. Optim.*, 19:377–398, 1998.
- [31] U. Tautenhahn. On the method of Lavrentiev regularization for nonlinear ill-posed problems. *Inverse Problems*, 18:191–207, 2002.
- [32] A. N. Tikhonov and V. B. Glasko. Use of the regularization method in non-linear problems. *USSR Comp. Math. Math. Phys.*, 5:93–107, 1965.