Johann Radon Institute for
Computational and Applied Mathematics
Austrian Academy of Sciences (ÖAW)

Approximate solution of nonlinear inverse problems by fixed-point iteration

S. Pereverzyev Jr., R. Pinnau, N. Siedow

RICAM-Report 2007-38

Approximate solution of nonlinear inverse problems by fixed-point iteration

Sergiy Pereverzyev Jr.*, René Pinnau† and Norbert Siedow‡

December, 2007

Abstract

In this paper we propose a derivative-free iterative method for the approximate solution of a nonlinear inverse problem $Fx = y$. In this method the iterations are defined as $Gx_{k+1} = Gx_k + (Sy - SFx_k)$, where G is an easily invertible operator and S is an operator from a data space to a solution space. We give general suggestions for the choice of operators G and S and show a practically relevant example of an inverse problem where such a method is succesfully applied. We carry out analysis of the proposed method for linear inverse problems. Using the recently introduced balancing principle we construct a stopping rule. Under reasonable assumptions, we show that this stopping rule leads to the regularization algorithm. Numerical results for a test example show its satisfactory behavior.

* *Technische Universität Kaiserslautern, Fachbereich Mathematik, Postfach 3049, 67653 Kaiserslautern, Germany;* and *Johann Radon Institute for Computational and Applied Mathematics (RICAM), Austrian Academy of Science, Altenbergerstrasse 69, 4040 Linz, Austria;* e-mail: `sergiy.pereverzyev@oeaw.ac.at`

† *Technische Universität Kaiserslautern, Fachbereich Mathematik, Postfach 3049, 67653 Kaiserslautern, Germany;* e-mail: `pinnau@mathematik.uni-kl.de`

‡ *Fraunhofer-Institut für Techno- und Wirtschaftsmathematik, Fraunhofer-Platz 1, 67663 Kaiserslautern, Germany;* e-mail: `siedow@itwm.fhg.de`

1 Introduction

A nonlinear inverse problem is often represented in the form of an operator equation

$$Fx = y \tag{1}$$

with a nonlinear *forward operator* F acting between a *solution space* X and a *data space* Y, which are assumed to be normed. A standard method for an approximate solution of (1) is Newton's method that is applicable to problems with Fréchet differentiable forward operators. In this method the following iterative procedure is considered:

$$x_{k+1} = x_k + (F'_{x_k})^{-1}(y - Fx_k), \tag{2}$$

where $F'_{x_k} : X \to Y$ is the Fréchet derivative of F at x_k.

In the context of *ill-posed* inverse problems special care must be taken in using (2). In such problems the inverse of the operator F as well as of its derivative F'_{x_k} is usually not continuous and is not defined on the whole data space Y. Then typically for some idealized data $y \in R(F)$ equation (1) has a unique solution $\hat{x} = \hat{x}(y)$. However for *noisy data* $y_\delta \in Y$ comming from real measurements equation (1) may have no solution or its solution may be far away from $\hat{x}$. In this case special techniques, called *regularization*, are needed [6].

There is substantial literature devoted to regularized versions of Newton's method, e.g. [1, 3, 7, 13, 5, 12]. A big disadvantage of (2) is the necessity to use Fréchet derivatives. First, its construction can be quite difficult; and secondly, its iterative update and inversion are usually very time consuming operations. Thus, the development of derivative-free methods is of interest. In this paper, we propose an iterative method that is based on rewriting equation (1) in the form of a fixed-point equation. Such a method does not use the Fréchet derivative and is easily implementable.

The paper is organized as follows. In Section 2 we introduce a general fixed-point iteration for nonlinear inverse problems and demonstrate its application to a practically relevant example. Then, in Section 3, we consider the behavior of the proposed fixed-point iterative process for linear inverse problems. Using the recently proposed balancing principle, we construct a stopping rule, which, as mentioned in Section 2, is needed when iterative processes are applied for noisy data. Under some assumptions motivated by numerical experience, we show that this stopping rule leads to the regularization algorithm. In Section 4 we present numerical results obtained by the proposed iterative process applied to the nonlinear inverse problem from Section 2 with noise-free data. We also test the numerical performance of the derived stopping rule on the example of a linear integral equation with noisy data. Finally, we finish with conclusions and an outlook.

2 Fixed-point iteration for nonlinear problems

In the case when the data space coincides with the solution space, equation (1) can be written in a form suggesting a solution method based on the fixed-point iteration

$$x_{k+1} = x_k + s(y - Fx_k), \quad s > 0. \tag{3}$$

If F is Lipschitz continuous and strongly monotone, then it is possible to choose s such that (3) converges to the solution of (1) (see, e.g., Zarantonello's theorem in [14, p. 54]).

For ill-posed inverse problems strong monotonicity of operator F can not be guaranteed. For just monotone operators a modification of (3) that guarantees convergence is proposed in [2]. However, satisfactory behavior of (3) is frequently observed in situations when no theoretical justification of the convergence is available. Moreover, the notable simplicity of (3) makes it very attractive for practical applications. Observe also that (3) can be modified for operator F acting between different function spaces. Namely, introducing a *scaling operator* $S : Y \to X$ one can use a generalized version of (3):

$$x_{k+1} = x_k + (Sy - SFx_k), \quad k = 0, 1, \ldots \tag{4}$$

For better convergence properties the introduction of an easily invertible operator $G : X \to X$ that modifies (4) as

$$Gx_{k+1} = Gx_k + (Sy - SFx_k), \quad k = 0, 1, \ldots \tag{5}$$

may be necessary. If $G = SF$, then independently of the initial guess x_0 the first iteration x_1 will already give us the solution of the original equation (1). But in most cases such G will not be easily invertible. Thus, a general suggestion might be $G \approx SF$. For this reason we call G the *approximating operator*.

Remark 1. We would like to point out that for linear approximating operators the iterative process (5) can be transformed to (4).

Example 1. In cooling processes of hot bodies, for example glass, the knowledge of the inside temperature is important. However, its direct measurement is usually very difficult, and therefore one tries to obtain it from other measurable quantities. In spectral pyrometry such a quantity is the radiation intensity measured on the surface of the cooling body. The problem of recovering the temperature from the radiation intensity can be represented in the form (1) [4]. We consider a particular case of this problem under the following assumptions:

- The body is a flat glass, and the temperature $x(z)$ is a function of thickness $z \in [0, 1]$.
- The surface corresponding to $z = 1$ is the black surface.

Then, temperature $x(z)$ satisfies the equation

$$\int_0^1 K(\lambda, z)B(x(z), \lambda)dz = y(\lambda), \quad \lambda \in [\lambda_1, \lambda_2],$$

where λ is the radiation wavelength, $[\lambda_1, \lambda_2]$ is the wavelength region where the radiation intensity is measured, $K(\lambda, z) = \exp(-\kappa(\lambda)z)$, $\kappa(\lambda) > 0$ is the material characteristic called the absorption coefficient, $B(x, \lambda) = [\lambda^5(\exp((x\lambda)^{-1}) - 1)]^{-1}$ is the scaled Planck function, $y(\lambda)$ is the function obtained from the outgoing radiation intensity at $z = 0$. For the description of modeling details we refer to [4].

A natural choice for the solution and data space is $X = L^2(0, 1)$, $Y = L^2(\lambda_1, \lambda_2)$. Then, the scalling operator can be choosen as

$$S : y(\lambda) \mapsto y(\varphi^{-1}(z)), \tag{6}$$

where $\varphi : [\lambda_1, \lambda_2] \to [0, 1]$ is some monotone function. We will consider the following form of φ:

$$\varphi(\lambda) = \left(\frac{\lambda_2 - \lambda}{\lambda_2 - \lambda_1}\right)^p, \quad p > 0. \tag{7}$$

Thus,

$$SFx(z') = \int_0^1 \tilde{K}(z', z)\tilde{B}(x(z), z')dz,$$

where $z' = \varphi(\lambda)$, $\tilde{K}(z', z) := K(\varphi^{-1}(z'), z)$, $\tilde{B}(x, z') := B(x, \varphi^{-1}(z'))$.

As approximating operator we propose to take

$$Gx(z') := \tilde{B}(x(z'), z') \int_0^1 \tilde{K}(z', z)dz. \tag{8}$$

It can be easily seen that

$$Gx(z') = SFx(z') \quad \text{for} \quad x(z) \equiv const.$$

Successfull performance of the numerical realization of (5) with the proposed scaling and approximating operators is demonstrated in Section 4.1.

Remark 2. The choice of φ in (7) can be motivated by the following considerations. The materials are usually more transparent for the waves with small wavelength λ ($\kappa(\lambda) \to 0$ as $\lambda \to 0$), and are almost opaque for the waves with large wavelength λ ($\kappa(\lambda) \to \infty$ as $\lambda \to \infty$). Therefore, one may expect to get the information about the temperature $x(z)$ in the remote points from the measurement cite (z close to 1) using the values of the intensity y for small λ; and vise versa, the information in the points nearest to the measurement cite (z close to 0) using the values of the intensity y for large λ. This suggests to consider the monotone functions in (6) satisfying $\varphi(\lambda_1) = 1$, $\varphi(\lambda_2) = 0$.

At this point it is worth to note that a theoretical justification of the convergence for a concrete problem is a very difficult analytical task. In any way, such a convergence can be only expected for the noise-free data, i.e. for $y \in R(F)$. However, as already mentioned, in practice only noisy data y_δ are available that in general do not belong to $R(F)$. Thus, apart from the question of convergence of (5) for noise-free data, another practically relevant question arises: what is the behavior of the iterative process (5) for noisy data?

Iterative processes for inverse and ill-posed problems ussually diverge for the noisy data. The same behavior can be expected for (5). Therefore, the design of reliable stopping rules is of big importance and interest. As the first step in understanding the behavior of (5) for noisy data, we are going to study the application of (5) to linear inverse problems.

3 Analysis for linear problems

Consider a linear inverse problem

$$Kx = y \tag{9}$$

with a linear forward operator $K : X \to Y$. We assume that K is injective and has nonclosed range $R(K) \subset Y$. Then, for every $y \in R(K)$ equation (9) has the unique solution $\hat{x} = \hat{x}(y)$. For linear problems it is reasonable to look for linear approximating and scaling operators. Then, according to Remark 1 it is sufficient to consider the iterative process (4).

Assumption 1. For every initial guess x_0 and for every $y \in R(K)$, the iterative process

$$x_{k+1} = x_k + (Sy - SKx_k), \quad k = 0, 1, \dots \tag{10}$$

converges to the corresponding solution $\hat{x}$ of (9).

Remark 3. The well-known Landweber iteration

$$x_{k+1} = x_k + K^*(y - Kx_k) \tag{11}$$

can be interpreted as a fixed-point iteration (10) with the scaling operator $S = K^*$. The usage of the adjoint operator was essential for proving the convergence of (11) because tools from spectral theory were used (see [6, Theorem 6.1]). Such tools cannot be employed for other scaling operators. The clear advantage of (10) is the possibility to use scaling operators which are much simpler than K^*, such as (6). However, in this case a verification of Assumption 1 becomes problematic.

Example 2. Consider the linear forward operator

$$Kx = \int_0^1 K(\lambda, z)x(z)dz, \quad \lambda \in [\lambda_1, \lambda_2]$$

with $K(\lambda, z) = \exp(-\kappa(\lambda)z)$ and $\kappa(\lambda)$ is defined in Example 1. It is a linear part of the operator from Example 1. Numerical experience shows that with the scaling operator defined by (6),(7) the discretized version of (10) is convergent with $p > 1$.

Assume that instead of Sy we are given noisy data Sy_δ such that

$$\|Sy - Sy_\delta\| \le \delta.$$

Therefore,

$$Sy - Sy_\delta = \delta\xi,$$

where $\xi \in X$ is such that $\|\xi\| \le 1$. We are interested in the behavior of the iterative process (10) with the noisy data

$$\begin{gathered} x_0^\delta = x_0, \\ x_{k+1}^\delta = x_k^\delta + (Sy_\delta - SKx_k^\delta), \quad k = 0, 1, \dots \end{gathered} \tag{12}$$

Consider the following estimate of the iteration error

$$\|\hat{x} - x_k^\delta\| \le \|\hat{x} - x_k\| + \|x_k - x_k^\delta\|. \tag{13}$$

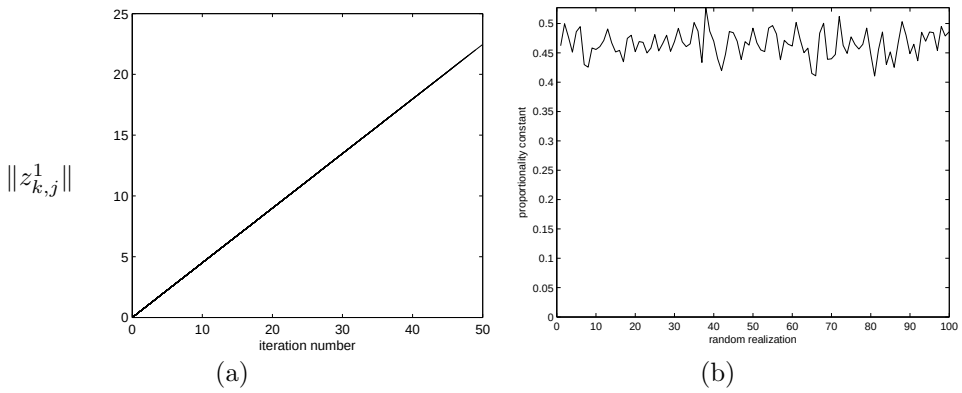


Figure 1: (a): typical behavior of $\|z^1_{k,j}\|$ for a fixed j; (b): variation of γ_j with j.

Estimates of this type are frequently used in the analysis of regularization methods. The first term $\|\hat{x} - x_k\|$ converges to zero, as $k \to \infty$, due to Assumption 1. Then, it is natural to assume that there exists a strictly monotone decreasing function $\psi : \mathbb{R}_+ \to \mathbb{R}_+$ with $\lim_{t \to +\infty} \psi(t) = 0$ such that

$$\|\hat{x} - x_k\| \leq \psi(k), \quad k = 0, 1, \ldots$$

Let us denote $z^\delta_k := x_k - x^\delta_k$. Then z^δ_k satisfies

$$z^\delta_0 = 0,$$
$$z^\delta_{k+1} = z^\delta_k + (\delta\xi - SKz^\delta_k), \quad k = 0, 1, \ldots$$

Hence, the behavior of z^δ_k is independent of the solution $\hat{x}$. Due to the linearity of SK it can be expressed as

$$z^\delta_k = \delta \sum_{i=0}^{k-1} (I - SK)^i \xi = \delta z^1_k.$$

Thus, it is sufficient to study z^1_k.

In general, the analytical derivation of an estimate for $\|z^1_k\|$ is as stubborn as a convergence justification for (10) (see Remark 3). Therefore, in practice it seems to be reasonable to study the behavior of $\|z^1_k\|$ numerically. To this end one looks at the behavior of

$$\|z^1_{k,j}\| = \left\| \sum_{i=0}^{k-1} (I - SK)^i \xi_j \right\|, \quad k = 1, 2, \ldots$$

for a random realization ξ_j and tries to find a consistency in its behavior.

We are going to illustrate this approach for the operator considered in Example 2. The values for λ_1, λ_2, p are taken from Table 1. Details of the random simulation of the scaled noise element are presented in Section A.2.

A typical behavior of $\|z^1_{k,j}\|$ for some random scaled noise element ξ_j is presented in Figure 1(a). One observes that

$$\|z^1_{k,j}\| \leq \gamma_j k,$$

where γ_j is a random variable. The variation of this variable with the number of random realization j is presented in Figure 1(b). This figure can be easily used for estimating an upper bound of γ_j and motivates the following assumption:

Assumption 2. There is a constant $\gamma > 0$ such that

$$\|z^1_k\| \leq \gamma k, \quad k = 0, 1, \ldots$$

Thus, under Assumptions 1 and 2 the triangle inequality (13) gives the error bound

$$\|\hat{x} - x^\delta_k\| \leq \psi(k) + \delta\gamma k \tag{14}$$

that can be used to design an appropriate stopping rule. In the regularization theory [6], a stopping rule is represented by a map $k : (\mathbb{R}_+, Y) \mapsto \mathbb{Z}_+$ that satisfies the condition

$$\|\hat{x} - x^{\delta}_{k(\delta,y_\delta)}\| \to 0 \quad \text{as} \quad \delta \to 0. \tag{15}$$

It can be easily seen that the choice of k balancing the terms in (14) as

$$k^{(\tau)}_* = k^{(\tau)}_*(\delta, y_\delta) := \min\{k : \ \psi(k) \le \tau\delta\gamma k\}, \quad \tau > 0 \tag{16}$$

satisfies (15). Moreover, it is possible to indicate the convergence rate that is independent of τ.

Denote $\theta_\psi(t) := \psi(t)/t$.

Lemma 1. *If the iterative process* (3) *is stopped by the stopping rule* (16)*, then the following estimate for the iteration error holds true:*

$$\|\hat{x} - x^{\delta}_{k^{\tau}_*}\| \le c\psi(\theta^{-1}_\psi(\delta)), \quad \delta < \delta_0, \tag{17}$$

where $c > 0$ is some constant independent of δ, and $\delta_0 > 0$ is an arbitrary but fixed upper bound for the noise level.

Proof. Let us introduce $t_* > 0$ and $t^{(\tau)}_* > 0$ satisfying

$$\psi(t_*) = \delta t_*,$$
$$\psi(t^{(\tau)}_*) = \delta(\tau\gamma)t^{(\tau)}_*.$$

Since $t^{(\tau)}_* \ge \theta^{-1}_\psi(\tau\delta_0\gamma) > 0$, there is a constant $c^{(\tau)}_* > 1$ independet of δ such that

$$t^{(\tau)}_* + 1 \le c^{(\tau)}_* t^{(\tau)}_*. \tag{18}$$

From (16) it follows that

$$t^{(\tau)}_* \le k^{(\tau)}_* \le t^{(\tau)}_* + 1. \tag{19}$$

Then, by (18),(19) and the monotonicity of ψ we get

$$\begin{aligned}\|\hat{x} - x^{\delta}_{k^{(\tau)}_*}\| &\le \psi(k^{(\tau)}_*) + (\delta\gamma)k^{(\tau)}_* \\ &\le \psi(t^{(\tau)}_*) + (\delta\gamma)(t^{(\tau)}_* + 1) \\ &\le \psi(t^{(\tau)}_*) + c^{(\tau)}_*(\delta\gamma)t^{(\tau)}_*.\end{aligned}$$

If $\tau\gamma < 1$, then $t_* < t^{(\tau)}_*$, and we have

$$\begin{aligned}\|\hat{x} - x^{\delta}_{k^{(\tau)}_*}\| &\le \left(1 + \frac{c^{(\tau)}_*}{\tau}\right)\psi(t^{(\tau)}_*) \\ &< \left(1 + \frac{c^{(\tau)}_*}{\tau}\right)\psi(t_*) = \left(1 + \frac{c^{(\tau)}_*}{\tau}\right)\psi(\theta^{-1}_\psi(\delta));\end{aligned}$$

otherwise, if $\tau\gamma \ge 1$, then $t_* \ge t^{(\tau)}_*$, and we have

$$\begin{aligned}\|\hat{x} - x^{\delta}_{k^{(\tau)}_*}\| &\le \left(\tau + c^{(\tau)}_*\right)\gamma\delta t^{(\tau)}_* \\ &\le \left(\tau + c^{(\tau)}_*\right)\gamma\delta t_* = \left(\tau + c^{(\tau)}_*\right)\gamma\psi(\theta^{-1}_\psi(\delta)).\end{aligned}$$

□

Remark 4. From the proof it follows that

$$\delta k^{(\tau)}_* \le c\psi(\theta^{-1}_\psi(\delta)), \quad \delta < \delta_0,$$

where $c > 0$ is some constant independent of δ, and $\delta_0 > 0$ is an upper bound for the noise level.

Note, the stopping rule (16) can seldomly be used in practice because the function ψ depends on the unknown solution and is rarely available. Using the recently introduced balancing principle [9, 10, 8], it is possible to design a stopping rule that does not use the function ψ but nevertheless leads to the convergence rate (17).

Theorem 1. *Let Assumptions 1 and 2 hold true and x_k^δ be the sequence generated by the iterative process* (10)*. Then, the iteration error produced by the following stopping rule:*

$$k_{\mathrm{bp}} = k_{\mathrm{bp}}(\delta, y_\delta) := \min\{k : \ \|x_k^\delta - x_m^\delta\| \le \tau_{\mathrm{bp}}\delta\gamma(k+m),\ (k+1) \le m \le k_{\max}(\delta)\}, \qquad (20)$$

where $\tau_{\mathrm{bp}} > 1$ and $k_{\max}(\delta) > k_*^{(\tau_{\mathrm{bp}}-1)}$, is estimated as

$$\|\hat{x} - x_{k_{\mathrm{bp}}}^\delta\| \le c\psi(\theta_\psi^{-1}(\delta)), \quad \delta < \delta_0,$$

where $c > 0$ is some constant independent of δ, and $\delta_0 > 0$ is an arbitrary but fixed upper bound for the noise level.

Proof. We will use ideas from the proof of Theorem 2.1 in [10]. First of all, observe that, using estimate (14) and definition of $k_*^{(\tau_{\mathrm{bp}}-1)}$ by (16), for all $k, m \ge k_*^{(\tau_{\mathrm{bp}}-1)}$ we have

$$\begin{aligned}\|x_k^\delta - x_m^\delta\| &\le \|\hat{x} - x_k^\delta\| + \|\hat{x} - x_m^\delta\| \\ &\le \psi(k) + \delta\gamma k + \psi(m) + \delta\gamma m \\ &\le \tau_{\mathrm{bp}}\delta\gamma(k+m).\end{aligned}$$

This implies

$$k_{\mathrm{bp}} \le k_*^{(\tau_{\mathrm{bp}}-1)}.$$

Finally,

$$\begin{aligned}\|\hat{x} - x_{k_{\mathrm{bp}}}^\delta\| &\le \left\|\hat{x} - x_{k_*^{(\tau_{\mathrm{bp}}-1)}}^\delta\right\| + \left\|x_{k_*^{(\tau_{\mathrm{bp}}-1)}}^\delta - x_{k_{\mathrm{bp}}}^\delta\right\| \\ &\le c_1\psi(\theta_\psi^{-1}(\delta)) + 2\tau_{\mathrm{bp}}\gamma\delta k_*^{(\tau_{\mathrm{bp}}-1)} \\ &\le (c_1 + 2\tau_{\mathrm{bp}}\gamma c_2)\psi(\theta_\psi^{-1}(\delta)),\end{aligned}$$

where $c_1 > 0$ is the constant from Lemma 1, and $c_2 > 0$ is the constant from Remark 4. □

Remark 5. We would like to point out that if an a priori bound

$$\psi(k) \le c, \quad k = 0, 1, \ldots$$

is available, then $k_{\max}(\delta)$ can be choosen as

$$k_{\max}(\delta) = \left\lceil \frac{c}{\delta\gamma(\tau_{\mathrm{bp}} - 1)} \right\rceil + 1.$$

In regularization theory there is another well-known stopping rule, the so-called *discrepancy principle*. According to this principle, the iterative process is stopped at

$$k_{\mathrm{dp}} := \min\{k : \ \|Sy_\delta - SKx_k^\delta\| \le \tau_{\mathrm{dp}}\delta\}, \qquad (21)$$

where $\tau_{\mathrm{dp}} > 0$ is a tuning parameter. We are going to compare the practical performance of the considered principles, however it should be pointed out that at the moment there is no rigorous justification of the discrepancy principle in the presented framework.

4 Numerical results

Details of the numerical realization of (5) can be found in Appendix A.1. Below we present results of application of the proposed method to operators from Examples 1 and 2.

	λ_1	λ_2	p	$\tau_{\rm bp}$	$\tau_{\rm dp}$
nonlinear problem	0.2	1.2	2	-	-
linear problem	0	5	8	1.1	2

Table 1: Values of the involved parameters considered in Section 4.

4.1 Application to the nonlinear inverse problem with noise-free data

We take the absorption coefficient as

$$\kappa(\lambda) = \frac{1}{8}\lambda^3 - \frac{9}{8}\lambda^2 + 3\lambda, \tag{22}$$

which is relevant to practical applications. The values of the parameters in (7) are taken from Table 1. In Figure 2 we show the results of numerical experiments with two solutions

$$\hat{x}_1(z) = 0.6 + 0.1z,$$
$$\hat{x}_2(z) = 0.7 - 0.1(z-1)^2,$$

as well as the iteration history. In both cases the initial guess is $x_0(z) \equiv 0.6$, and exact values of the right hand side y are used. One observes rather fast and accurate reconstruction.

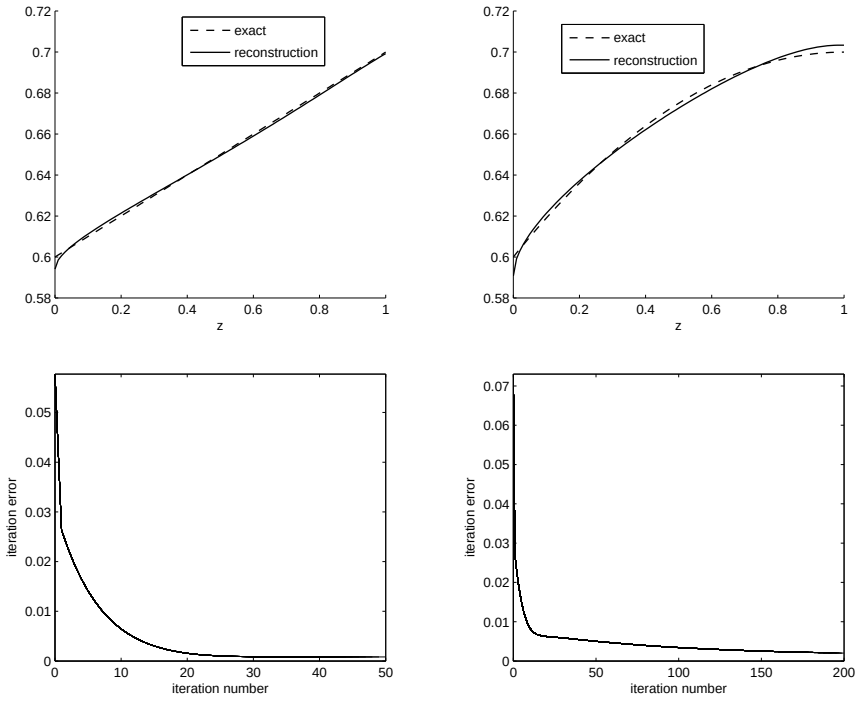


Figure 2: Experiments with two different solutions.

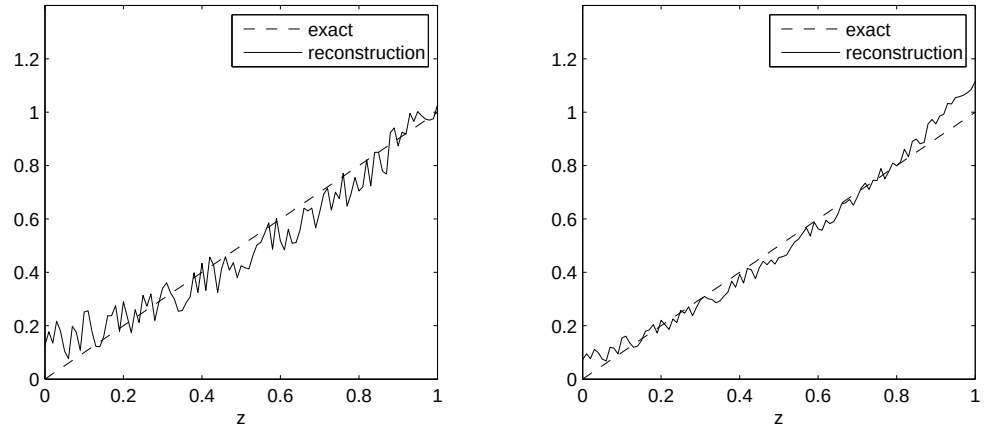
4.2 Application to the linear inverse problem with noisy data

Consider now the linear forward operator from Example 2. The function $\kappa(\lambda)$ is taken as in (22). We test the numerical performance of the iterative process (3) with the stopping rules (20) and (21). The values of the involved parameters are given in Table 1.

We present the numerical results for the following exact solutions:

$$\hat{x}_1(z) = z, \tag{23}$$
$$\hat{x}_2(z) = 4(z - z^2). \tag{24}$$

$\delta = 0.01$ $\delta = 0.001$



	mean value	standard deviation
BP	6.778e-02	3.079e-03
DP	7.170e-02	2.931e-03
BEST	6.544e-02	3.834e-03

	mean value	standard deviation
BP	3.977e-02	1.925e-03
DP	4.006e-02	1.389e-03
BEST	3.897e-02	2.191e-03

Figure 3: Numerical results for profile (23). Typical reconstructions obtained by the balancing principle (20). Mean values and standard deviations of the reconstruction errors of different stopping rules for 100 independent random realizations of the noisy data (BP – balancing principle (20), DP – discrepancy principle (21), BEST – stopping rule with the smallest error (25)).

The following noise levels are taken: $0.01, 0.001$.

Typical reconstructions obtained with the balancing principle (20) are shown in Figures 3 and 4. Comparison of the mean values and standard deviations of the reconstruction errors produced by the balancing principle, discrepancy principle, and by the stopping rule that gives the smallest error, i.e.

$$k_{\text{best}} := \text{argmin}\{\|\hat{x} - x_k\|,\ k = 0, \ldots, k_{\max}\}, \tag{25}$$

for 100 independent realizations of the noisy data are presented in the tables of Figures 3 and 4.

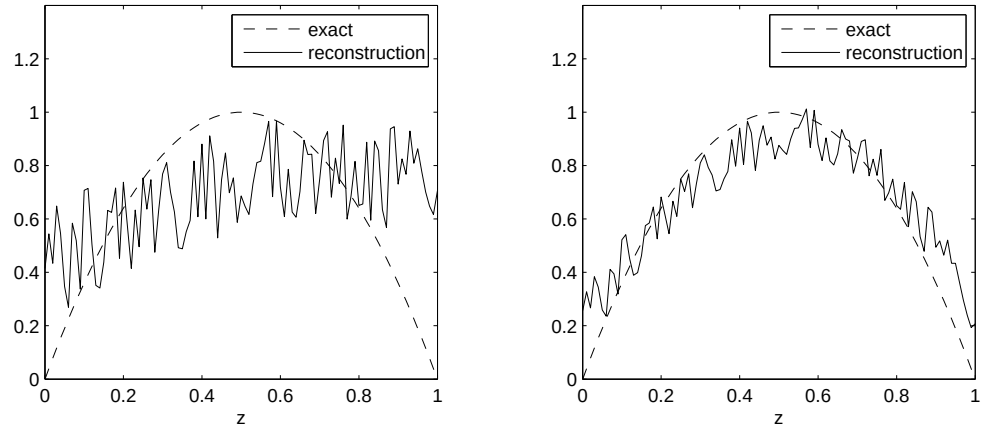
Both stopping rules (20) and (21) show satisfactory behavior, although a theoretical justification is at the moment only available for the balancing principle (20).

5 Conclusions and outlook

In this paper we introduced a general derivative-free fixed-point iterative method for the approximate solution of nonlinear inverse problems. Its regularization properties were studied for linear problems. Using the recently introduced balancing principle, we constructed a stopping rule and, under assumptions derived from the numerical experience, proved that it leads to the regularization method. Numerical results for a test example show its satisfactory behavior. We would like to note that the idea of representing the nonlinear inverse problem as a fixed-point equation in order to avoid the usage of the Fréchet derivative was also used in [11]. The resulting iterative process in [11] required to solve an ill-posed equation on each iteration, while each iteration step in the method proposed here is well-posed and much easily realizable. However, the overall iterative process in [11] was stable, i.e. the iteration error did not explode with the increase of the iteration number. Therefore, in contrast to the method considered here, the choice of the stopping rule was not crucial. The comparison of the methods with respect to the reconstruction quality and computing time is of interest. Other issues for the future research can be indicated:

- theoretical justification of the discrepancy principle for the introduced iterative scheme;
- more extensive numerical and analytical comparison of the balancing principle with the discrepancy principle;
- theoretical and numerical study of the proposed iterative method for nonlinear problems.

$\delta = 0.01$ $\delta = 0.001$



	mean value	standard deviation
BP	2.727e-01	1.440e-02
DP	3.222e-01	1.025e-02
BEST	2.667e-01	1.554e-02

	mean value	standard deviation
BP	1.141e-01	4.900e-03
DP	1.058e-01	5.279e-03
BEST	0.958e-01	4.620e-03

Figure 4: Numerical results for profile (24). Typical reconstructions obtained by the balancing principle (20). Mean values and standard deviations of the reconstruction errors of different stopping rules for 100 independent random realizations of the noisy data (BP – balancing principle (20), DP – discrepancy principle (21), BEST – stopping rule with the smallest error (25)).

Acknowledgements

Authors acknowledge support from the *DFG Schwerpunktprogramm 1253* project PI 408/3-1. The authors also would like to thank Professor Rainer Kress (Universität Göttingen) for the interest in this work and helpful comments.

A Appendix

A.1 Numerical approximation of the forward operators

The operators considered in Examples 1 and 2 are particular cases of a general nonlinear integral operator

$$F(x(z)) = \int_0^1 L(z', z, x(z))dz.$$

We realize their numerical approximations as follows:

- the domain $[0, 1]$ for variables z,z' is discretized by an equidistant grid with n points;
- the integral $\int_0^1 L(z', z, x(z))dz$ at each grid point z' is approximated by trapezoidal rule.

All presented numerical results were done with $n = 101$.

A.2 Generation of the noisy data

The noise free data were obtained by computing on the grid with $n = 1001$. Noisy data were obtained by projecting noise free data on the computational grid and then adding noise element $\delta\xi$. Below we describe the random realization of the scaled noise element.

First, we consider a piecewise linear spline ξ' with knots on the computational grid and with values in the knots to be the random variable uniformly distributed on $[-1, 1]$. It is clear that $\|\xi'\| \leq 1$. In Figure 5 the norms of 100 random realizations of ξ' are depicted, from which we may conclude that the norm of ξ' can be estimated by a smaller constant. In our numerical experiments we assume that

$$\|\xi'\| \leq 0.54 =: \tau'.$$

Then, the scaled noise element is

$$\xi = \frac{\xi'}{\tau'}.$$

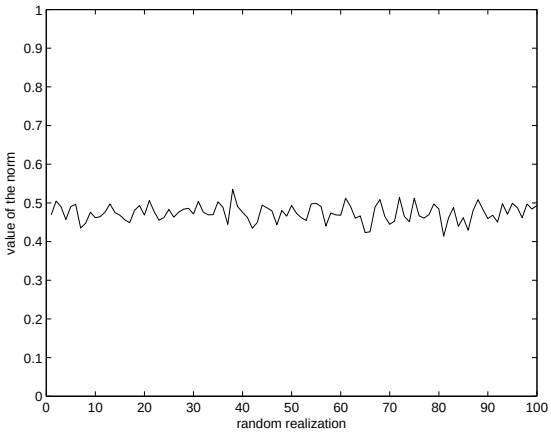


Figure 5: Variation of the norm $\|\xi'\|$.

References

[1] A. B. Bakushinskij. The problem of the convergence of the iteratively regularized Gauss-Newton method. *Comput. Math. Math. Phys.*, 32:1353–1359, 1992.

[2] A. B. Bakushinsky and A. B. Smirnova. A posteriori stopping rule for regularized fixed-point iterations. *Journal of Nonlinear Analysis Series A: Theory, Methods and Applications*, 64:1255–1261, 2006.

[3] 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.

[4] M. Brinkmann, N. Siedow, and T. Korb. Remote spectral temperature profile sensing. In D. Krause and H. Loch, editors, *Mathematical Simulation in Glass Technology*, pages 262–286. Springer-Verlag, 2002.

[5] P. Deuflhard, H. W. Engl, and O. Scherzer. A convergence analysis of iterative methods for the solution of nonlinear ill-posed problems under affinely invariant conditions. *Inverse Problems*, 14:1081–1106, 1998.

[6] H. W. Engl, M. Hanke, and A. Neubauer. *Regularization of inverse problems*. Kluwer Academic Publishers, Dordrecht, 1996.

[7] B. Kaltenbacher. Some Newton-type methods for the regularization of nonlinear ill-posed problems. *Inverse Problems*, 13:729–753, 1997.

[8] R. D. Lazarov, S. Lu, and S. Pereverzyev. On the balancing principle for some problems of numerical analysis. *Numerische Mathematik*, 106:659 – 689, 2007.

[9] P. Mathe and S. V. Pereverzev. Geometry of linear ill-posed problems in variable Hilbert scales. *Inverse Problems*, 19:789–803, 2003.

[10] S. Pereverzev and E. Schock. On the adaptive selection of the parameter in regularization of ill-posed problems. *SIAM J. Numer. Anal.*, 43:2060–2076, 2005.

[11] S. S. Pereverzyev, R. Pinnau, and N. Siedow. Regularized fixed-point iterations for nonlinear inverse problems. *Inverse Problems*, 22:1–22, 2006.

[12] A. Rieder. On convergence rates of inexact Newton regularizations. *Numer. Math.*, 88:347–365, 2001.

[13] U. Tautenhahn. On a general regularization scheme for nonlinear ill-posed problems. *Inverse Problems*, 13:1427–1437, 1997.

[14] G. Teschl. Nonlinear functional analysis. lecture notes at http://www.mat.univie.ac.at/ gerald/ftp/book-nlfa/index.html, 2005.