Zentrum für Technomathematik

Fachbereich 3 – Mathematik und Informatik

TIGRA – an iterative algorithm for regularizing nonlinear ill–posed problems

Ronny Ramlau

Report 02–07

Berichte aus der Technomathematik

Report 02–07 **August 2002**

TIGRA–an iterative algorithm for regularizing nonlinear ill-posed problems

Ronny Ramlau
University of Bremen, Germany

August 21, 2002

Abstract

We report on a new iterative method for regularizing a nonlinear operator equation in Hilbert spaces. The proposed TIGRA-algorithm is a combination of TIKhonov–regularization and a GRAdient method for minimizing the Tikhonov–functional. Under the assumptions that the operator F is twice continuous Fréchet–differentiable with Lipschitz–continuous first derivative and that the solution of the equation $F(x) = y$ fulfills a smoothness condition we will give a convergence rate result. Finally we present some applications and a numerical result for the reconstruction of the activity function in Single Photon Emission Computed Tomography (SPECT).

AMS Classification. 65J15, 65J20, 65J22, 44A12

1 Introduction

This paper is dedicated to the stable solution of a nonlinear ill–posed operator equation

$$F(x) = y \ , \tag{1}$$

where $F: \ X \to Y$ is a map between real Hilbert spaces X, Y. If only noisy data y^δ with

$$\|y^\delta - y\| \le \delta \tag{2}$$

are available, then the problem of solving (1) has to be *regularized.* Due to the importance for technical applications, many of the known regularization methods for linear operator equations have been generalized to nonlinear equations during the last decade. Roughly speaking, the developed methods can be separated into two classes: Tikhonov–regularization and iteration methods. Tikhonov–regularization might be the best known method. As an approximation to a solution of (1), one takes a minimizer x_α^δ of the Tikhonov–functional

$$\Phi_\alpha(x) = \|y^\delta - F(x)\|^2 + \alpha\|x - \bar{x}\|^2 \ , \tag{3}$$

$\alpha > 0$. The function $\bar{x}$ plays the role of a selection criterion, x_α^δ is an approximation to a solution of (1) with minimal distance to $\bar{x}$. The theoretical results for Tikhonov–regularization are encouraging: Under rather weak conditions to the nonlinear operator F it can be shown that, for properly chosen regularization parameter α, the minimizing function x_α^δ is a good approximation to a solution x_* with minimal distance to $\bar{x}$ (we will refer to such a solution as $\bar{x}$–minimum–norm solution). Under a smoothness condition to the solution, convergence rates can be shown [9]. However, there are some open problems. In most of the results it is assumed that the minimizing element x_α^δ of (3) is *known*. For a numerical realization, x_α^δ (or a sufficient good approximation) has to be computed. For a *linear* operator A, the minimizer of (3) is unique and can be computed by

$$x_\alpha^\delta = (A^*A + \alpha I)^{-1} A^* y^\delta \ .$$

For a nonlinear operator F, (3) might have several (local) minima. Using an optimization routine for minimizing the Tikhonov–functional one might end up with only a local minimizer of (3), and therefore with a bad approximation to the solution x_*. Another problem is the selection of the regularization parameter α. It is well known that a posteriori parameter choice rules like Morozov's discrepancy principle in general yield better results than a priori rules. But for general nonlinear operators F a parameter fulfilling Morozov's discrepancy principle might not exist (for more details, cf. Sections 2 and 6).
In contrary to Tikhonov–regularization, iteration methods produce an approximation to the solution within every iteration step. Due to the ill–posedness of the operator, the iteration has to be terminated early enough, and the termination index of the iteration is then the regularization parameter. Several iteration methods for nonlinear operators were under investigation during the last years, e.g. Landweber methods [13, 23], Levenberg–Marquardt methods [11], Gauss–Newton [1, 2], conjugate gradient [12] and other Newton–like methods [15]. The computation of a new iterate is mostly not difficult to perform. For Landweber's method, one has to evaluate the operator F and the adjoint of the Fréchet-derivative of F. For all other methods, one has to solve a linear equation additionally.
The drawback is that the analysis of these methods appears to be more difficult than for Tikhonov–regularization. In order to show regularization properties, convergence rates or to introduce stopping rules for the iteration, one has to impose rather stringent conditions on the operator and its Fréchet derivative. For example, the convergence analysis for Landweber iteration was given in [13] under the condition

$$\|F(\tilde{x}) - F(x) - F'(x)(x - \tilde{x})\| \leq \eta \|F(\tilde{x}) - F(x)\| \ , \qquad \eta < 1/2 \ , \tag{4}$$

and in [5] a Newton–Mysovskii condition,

$$\|(F'(x) - F'(x_*))F'(x_*)^\#\| = O(\|x - x_*\|) \ , \tag{5}$$

($F'(x)^\#$ denotes the left inverse of $F'(x)$) was used to analyze Landweber's method and the iteratively regularized Gauss–Newton method. However, for practical applications, eg. from medical imaging, it is difficult to prove the required estimates, and it would be therefore of interest to develop iteration methods for which convergence results can be shown under weak restrictions to the operator.
By reviewing the analysis of the above mentioned iteration methods one gets the idea that conditions like (4) have to be introduced mainly due to the ill–posedness of the operator F. If one considers Landweber iteration, then this is a steepest descent method for minimizing the

residual $R(x) = \|y^\delta - F(x)\|$. It is well known [17] that the steepest descent method converges to the unique minimizer of the functional if $R(x)$ is convex. But even for linear operators this condition is violated in case of an ill–posed operator. On the other hand, we have shown in [25] that the Tikhonov–functional (3) is locally convex for bilinear operators, and that the steepest descent method can be used to find a minimizer of $\Phi_\alpha(x)$. Thus the main goal of this paper will be the development of an iterative algorithm which combines Tikhonov–regularization with the steepest descent method. The resulting TIGRA–method (Tikhonov–Gradient–method) will be defined by

$$x^\delta_{k+1} = x^\delta_k + \beta_k(F'(x^\delta_k)^*(y^\delta - F(x^\delta_k)) + \alpha_k(x^\delta_k - \bar{x})) \ . \tag{6}$$

Here, β_k denotes a scaling parameter. We might remark that in contrary to the steepest descent method for minimizing Φ_α, the regularization parameter α_k will change during the iteration. As a stopping rule for the iteration we will use a Morozov–like discrepancy principle. We might remark that Scherzer [28] considered a similar iteration procedure under the much stronger condition (4) to the operator F. He named it a modified Landweber method, but as our approach for the analysis of the iteration as well as for the choice of the parameters is totally different from [28] and is based on the interaction of Tikhonov–regularization with the gradient method, we feel that Tikhonov–Gradient method is an appropriate name for the iteration.

Throughout this paper we will assume that the following conditions hold:

1. F is twice Fréchet–differentiable with continuous second derivative ,

2. The first derivative is Lipschitz–continuous,

$$\|F'(x_1) - F'(x_2)\| \le L\|x_1 - x_2\| \ . \tag{7}$$

3. There exists $\omega \in Y$ with

$$x_* - \bar{x} = F'(x_*)^*\omega \ , \tag{8}$$

4. and

$$\|\omega\| \le \varrho \ \text{ and } L\varrho \le 0.278 \ . \tag{9}$$

We will show that the TIGRA–algorithm is a stable method under the above conditions and will give a convergence rate result. Moreover, we might remark that our conditions are only slightly more restrictive than the conditions for the classical convergence rate result for Tikhonov–regularization given in [10]. (cf. Section 2).

The structure of the paper will be as follows: In Section 2 we will collect some well known results for Tikhonov–regularization for nonlinear operators. Section 3 contains results from operator calculus for nonlinear operators. In Section 4 it will be shown that the Tikhonov–functional has a convexity property in the neighborhood of a global minimizer of the functional. A convergence analysis for the steepest descent algorithm for minimizing the Tikhonov–functional is given in Section 5 . It is shown that the algorithm converges to a global minimizer of the Tikhonov–functional if a starting value for the iteration within the above mentioned neighborhood is known. Section 6 is dedicated to the investigation of the continuity of the mapping $\alpha \to x^\delta_\alpha$, where x^δ_α denotes a global minimizer of Φ_α. The main result is that this mapping is continuous under the above given assumptions, and that especially a regularization parameter α exists such that Morozov's discrepancy principle holds. The TIGRA-algorithm is introduced in Section 7, and a convergence rate result is proven. Finally, in Section 8, we will give some applications.

2 Some results on Tikhonov–regularization

In this section we will collect results on Tikhonov–regularization which will be used later on. In 1989 Engl, Kunisch and Neubauer [10] gave the first convergence results for Tikhonov–regularization. If a Fréchet differentiable operator F fulfills (7) and a smoothness condition (8) with $L\|\omega\| < 1$ is given, then they could show for an a priori parameter choice $\alpha \sim \delta$ the error estimate

$$\|x_\alpha^\delta - x_*\| = O(\sqrt{\delta}) \ . \tag{10}$$

For our purpose, we will need an intermediate result of the proof for this Theorem. The authors show that the following estimates hold:

$$\|y^\delta - F(x_\alpha^\delta)\| \ \leq \ \delta + 2\alpha\|\omega\| \ , \tag{11}$$

$$\|x_\alpha^\delta - x_*\| \ \leq \ \frac{\delta + \alpha\|\omega\|}{\sqrt{\alpha}\sqrt{1 - L\|\omega\|}} \ . \tag{12}$$

The proof of the Theorem and the above estimates can be found additionally in [9], pp. 245. We might remark that for these estimates no information about the parameter choice was used. Setting $\alpha = \delta/\eta$, we obtain

$$\|x_\alpha^\delta - x_*\| \leq \frac{\sqrt{\eta}(1 + \|\omega\|/\eta)}{\sqrt{1 - L\|\omega\|}}\delta^{1/2} =: c(\eta)\delta^{1/2} \ .$$

The function $c(\eta)$ has a minimum at $\eta = \|\omega\|$, and so the best a priori parameter choice would be $\alpha_{opt} = \delta/\|\omega\|$. Of course, α_{opt} is usually not known, but ϱ with $\|\omega\| \leq \varrho$ might be available and so we could use $\tilde{\alpha}_{opt} = \delta/\varrho \leq \alpha_{opt}$ instead. It is well known that a too small regularization parameter might cause a higher instability, and it seems therefore reasonable to allow only regularization parameters such that $\tilde{\alpha}_{opt} \leq \alpha = \delta/\eta$, i.e. to choose $\eta \leq \varrho$.
It is well known that a posteriori parameter choice rules lead in general to better results than a priori rules. One of the most used rule is Morozov's discrepancy principle, where a regularization parameter α is chosen such that

$$\|y^\delta - F(x_\alpha^\delta)\| = c\delta \ , \qquad c \geq 1, \tag{13}$$

holds. If such a parameter exist, then it is easy to show that x_α^δ fulfills an error estimate (10). In [24] a modified discrepancy principle was used. There, the regularization parameter was determined such that

$$\delta \leq \|y^\delta - F(x_\alpha^\delta)\| \leq c\delta \ , \qquad c \geq 1, \tag{14}$$

holds, which is more realistic for practical applications. If such a parameter exists, then the error estimate

$$\|x_\alpha^\delta - x_*\| \leq \left(\frac{2(1+c)\|\omega\|}{1 - L\|\omega\|}\right)^{1/2} \delta^{1/2} \tag{15}$$

holds, cf. [24]. Additionally, some simple conditions under which a parameter with (14) exists were given in this paper.
Besides the question of the proper choice of the regularization parameter, for arbitrary nonlinear operators it is an open question how the minimizer of the functional can be computed. In [25] we have shown that the function

$$\varphi_{\alpha,h}(t) = \Phi_\alpha(x_\alpha^\delta + th) \qquad t \in \mathbb{R},\ h \in X$$

for a *bilinear* operator F is a convex function for all t, h with $\|th\| \leq r(\alpha)$ and $r(\alpha) = O(\alpha + \sqrt{\alpha})$. As a consequence the steepest descent method could be used to minimize Φ_α for this class of operators. In Sections 4 and 5 this result will be generalized to twice Fréchet differentiable operators with Lipschitz–continuous derivative.

3 Some results from operator calculus

The analysis of Tikhonov–regularization and the TIGRA–algorithm in the following sections will be based mainly on the Taylor–series for F and especially on the fact, that for twice continuous Fréchet–differentiable operators the second order remainder of the series can be expressed in integral form, i.e. we have

$$F(x+h) = F(x) + F'(x)h + \frac{1}{2}\int_0^1 (1-\tau)F''(x+\tau h)(h,h)\,d\tau \tag{16}$$

Due to the Lipschitz–continuity (7), the second derivative can globally be estimated by

$$\|F''(x)(h,h)\| \leq L\|h\|^2 \ . \tag{17}$$

Moreover, we obtain for the first Fréchet–derivative

$$F'(x+h) = F'(x) + \int_0^1 F''(x+\tau h)(h,\cdot)\,d\tau \tag{18}$$

In the following section we have to use some properties of the integrals given in (16) and (18), and so we might esp. recall the definition of an integral of a vector function of one real variable: If $f : [a,b] \to Y$ is a mapping into a Banach–space Y and $-\infty < a < b < \infty$, then we can define the partial sums

$$S_{Z_n} = \sum_{i=1}^{n} f(\bar{t}_i)(t_i - t_{i-1}), \qquad t_{i-1} \leq \bar{t}_i \leq t_i \ ,$$

where Z_n is a partition $a = t_0 < t_1 < \cdots < t_n = b$ of $[a,b]$ and $\Delta Z_n = \max\{t_i - t_{i-1}\}$ is the mesh of the partition. As for the classical Riemann integral, we define

$$\int_a^b f(t)\,dt = \lim_{n\to\infty} S_{Z_n}, \tag{19}$$

if such a common limit exists for all sequences of partitions (Z_n) with $\Delta Z_n \to 0$ as $k \to \infty$. It is well known that the integral exists if f is continuous on $[a, b]$. With respect to (18), we can define for fixed h the operator $B_h : X \to Y$ by

$$B_h k := \int_0^1 F''(x + \tau h)(h, k)\, d\tau \tag{20}$$

and observe that B_h is a linear operator in k: Indeed, the partial sums for the above integral,

$$S_{Z_n}(k) = \sum_{i=1}^{n} F''(x + \bar{\tau}_i h)(h, k)(\tau_i - \tau_{i-1}),$$

are linear in k and thus the limit, B_h, is linear. Thus we obtain from (18)

$$F'(x + h) = F'(x) + B_h \tag{21}$$

and especially

$$F'(x + h)^* = F'(x)^* + B_h^* \ , \tag{22}$$

which we will need later. Another frequently used property of the integral is

$$\int_0^1 f(a\tau)\, d\tau = \frac{1}{a} \int_0^a f(\tau)\, d\tau \ . \tag{23}$$

Indeed, by setting $\kappa_i = a\tau_i$, $i = 1, \ldots n$, we find $\kappa_0 = 0$, $\kappa_n = a$ and

$$\sum_{i=1}^{n} f(a\bar{\tau}_i)(\tau_i - \tau_{i-1}) \quad = \quad \frac{1}{a} \sum_{i=1}^{n} f(\bar{\kappa}_i)(\kappa_i - \kappa_{i-1}) \ ,$$

where the right hand side of the last equation is a partition of the integral on the right hand side of (23). As for Riemann integrals, we get the Fundamental Theorem of Calculus: If

$$G(t) = \int_a^t g(\tau)\, d\tau \ , \tag{24}$$

then G is differentiable with

$$G'(t) = g(t) \ . \tag{25}$$

For a proof, we refer to [32], Part I, p. 77.

4 A convexity property of the Tikhonov–functional

Tikhonov–regularization requires the computation of a global minimizer of the functional

$$\Phi_\alpha(x) := \|y^\delta - F(x)\|^2 + \|\bar{x} - x\|^2 \ . \tag{26}$$

For nonlinear operators F it might be difficult to compute a global minimizer of the Tikhonov–functional. There might exist several global minima and even local minima. In such a situation, most of the known algorithms will only manage to converge to a local minimum. Our aim is to use the steepest descent method for minimizing Φ_α. It is well known that steepest descent will converge to (the global) minimizer of a functional, if it is convex. In [25] it was shown that this condition can be relaxed for the Tikhonov–functional with bilinear operator F. In the following, this result will be generalized for arbitrary nonlinear operators F with Lipschitz–continuous first derivative and continuous second derivative.
For what comes we will set

$$C_x(t)(h,h) := \frac{1}{2}\int_0^1 (1-\tau)F''(x+t\tau h)(h,h)\,d\tau \ . \tag{27}$$

We obtain with (16)

$$F(x+th) = F(x) + t\cdot F'(x)h + t^2\cdot C_x(t)(h,h) \tag{28}$$

and

$$\begin{aligned}\Phi_\alpha(x+th) \;=\;& \Phi_\alpha(x) - 2t\left(\langle y^\delta - F(x), F'(x)h\rangle - \alpha\langle x-\bar{x}, h\rangle\right)\\ &+t^2(\|F'(x)h\|^2 - 2\langle y^\delta - F(x), C_x(t)(h,h)\rangle + \alpha\|h\|^2)\\ &+2t^3\langle F'(x)h, C_x(t)(h,h)\rangle\\ &+t^4\|C_x(t)(h,h)\|^2 \ .\end{aligned} \tag{29}$$

Proposition 4.1 *Let $C_x(t)(h,h)$ be defined as in (27), and define $g(t):\mathbb{R}^+\to Y$ by*

$$g(t) := t^2 C_x(t)(h,h) \ . \tag{30}$$

Then $g(t)$ is twice differentiable and the following properties hold:

i) $$g(t) = \frac{1}{2}\int_0^t (t-\tau)F''(x+\tau h)(h,h)\,d\tau \ , \tag{31}$$

ii) $$g'(t) = \frac{1}{2}\int_0^t F''(x+\tau h)(h,h)\,d\tau = \frac{t}{2}\int_0^1 F''(x+t\tau h)(h,h)\,d\tau \ , \tag{32}$$

iii) $$g''(t) = \frac{1}{2}F''(x+th) \ , \tag{33}$$

and

iv) $$\left[\|g(t)\|^2\right]'' = 2(\langle g''(t), g(t)\rangle + \|g'(t)\|^2) \ . \tag{34}$$

Proof:
i) By using (23) with $f(t\tau) = (t-t\tau)F''(x+t\tau h)(h,h)$.
ii) By using (31), the product rule, (25) and (23).

iii) Follows from (32) and (25).
iv) For differentiable mappings $u(t), v(t) : \mathbb{R} \to Y$ we have

$$\langle u(t), v(t)\rangle' = \langle u(t), v'(t)\rangle + \langle u'(t), v(t)\rangle \ , \tag{35}$$

and applying this formula twice gives (34).

□

In the following we would like to investigate the area of convexity of the function

$$\varphi_{\alpha,h}(t) := \Phi_\alpha(x_\alpha^\delta + th) \ , \qquad t \in \mathbb{R},\ \|h\| = 1 \ . \tag{36}$$

where x_α^δ denotes a (global) minimizer of the functional Φ_α. We have shown in [25] that $\varphi_{\alpha,h}(t)$ with a *bilinear* operator is convex as long as $|t| \leq \tilde{r}(\varrho)(\alpha + \sqrt{\alpha})$ holds, where $\tilde{r}$ is a positive function. For twice Fréchet-differentiable nonlinear operators we will only be able to show that $\varphi_{\alpha,h}(t)$ is convex for $|t| \leq r(\varrho)\min\{\alpha, \sqrt{\alpha}\}$.

Proposition 4.2 *Let $C_{x_\alpha^\delta}(t)(h,h)$ be as in (27) and $\tilde{C}_{x_\alpha^\delta}(t)(h,h)$ be defined by*

$$\tilde{C}_{x_\alpha^\delta}(t)(h,h) := \frac{1}{2}\int_0^1 F''(x_\alpha^\delta + t\tau h)(h,h)\,d\tau \ . \tag{37}$$

Then we obtain for the second derivative of $\varphi_{\alpha,h}(t)$ with $\|h\| = 1$

$$\begin{aligned} \varphi''_{\alpha,h}(t) \ = \ & 2\|F'(x_\alpha^\delta)h\|^2 - \langle y^\delta - F(x_\alpha^\delta), F''(x_\alpha^\delta + th)(h,h)\rangle + 2\alpha \\ & + t\langle F'(x_\alpha^\delta)h, 4\tilde{C}_{x_\alpha^\delta}(t)(h,h) + F''(x_\alpha^\delta + th)(h,h)\rangle \\ & + t^2\langle F''(x_\alpha^\delta + th)(h,h), C_{x_\alpha^\delta}(t)(h,h)\rangle \\ & + 2t^2\|\tilde{C}_{x_\alpha^\delta}(t)(h,h)\|^2 \ . \end{aligned} \tag{38}$$

Proof:
x_α^δ is a minimizer of Φ_α, and thus

$$\Phi_\alpha{}'(x_\alpha^\delta) = -2(\langle y^\delta - F(x), F'(x)h\rangle + \alpha\langle x - \bar{x}, h\rangle) = 0$$

(cf. (29)). Defining $g(t)$ as in (30) and x replaced by x_α^δ, $\varphi_{\alpha,h}(t)$ can be written by (29) as

$$\begin{aligned} \varphi_{\alpha,h}(t) \ = \ & \varphi_{\alpha,h}(0) + t^2(\|F'(x_\alpha^\delta)h\|^2 + \alpha) - 2\langle y^\delta - F(x_\alpha^\delta), g(t)\rangle \\ & + 2\langle t \cdot F'(x_\alpha^\delta)h, g(t)\rangle + \|g(t)\|^2 \ . \end{aligned} \tag{39}$$

We have

$$\begin{aligned} (t^2(\|F'(x_\alpha^\delta)h\|^2 + \alpha))'' \ &= \ 2(\|F'(x_\alpha^\delta)h\|^2 + \alpha) \ , \\ \langle y^\delta - F(x_\alpha^\delta), g(t)\rangle'' \ &= \ \langle y^\delta - F(x_\alpha^\delta), g''(t)\rangle \ . \end{aligned}$$

Using (35) with $u(t) = t \cdot F'(x_\alpha^\delta)h$, $v(t) = g(t)$ we get

$$\langle u, v\rangle'' = \langle u'', v\rangle + \langle u, v''\rangle + 2\langle u', v'\rangle \ ,$$

and, because of $u'(t) = F'(x_\alpha^\delta)h$, $u''(t) = 0$

$$\langle t \cdot F'(x_\alpha^\delta)h, g(t)\rangle'' = t\langle F'(x_\alpha^\delta)h, g''(t)\rangle + 2\langle F'(x_\alpha^\delta)h, g'(t)\rangle \ .$$

Thus we have found

$$\varphi''_{\alpha,h}(t) = 2(\|F'(x_\alpha^\delta)h\|^2+\alpha)-\langle y^\delta-F(x_\alpha^\delta), g''(t)\rangle+2t\langle F'(x_\alpha^\delta)h, g''(t)\rangle+4\langle F'(x_\alpha^\delta)h, g'(t)\rangle+\left[\|g(t)\|^2\right]'' \ ,$$

and using (31)–(34), (37) yields (38).

□

Proposition 4.3 *Let $\delta = \eta\alpha$ with $\eta \le \varrho$. The norm of $F'(x_\alpha^\delta)$ can be estimated by*

$$\|F'(x_\alpha^\delta)\| \le \max\left\{L\|y^\delta - F(\bar{x})\| + \|F'(\bar{x})\|, \frac{2L\varrho}{\sqrt{1-L\varrho}} + \left(\frac{1}{1-L\varrho}+1\right)\|F'(\bar{x})\|\right\} =: K \ . \quad (40)$$

Proof:
It is for $\alpha \ge 1$

$$\begin{aligned} \|F'(x_\alpha^\delta)\| &\le \|F'(x_\alpha^\delta) - F'(\bar{x})\| + \|F'(\bar{x})\| \\ &\overset{(7)}{\le} L\|x_\alpha^\delta - \bar{x}\| + \|F'(\bar{x})\| \ , \end{aligned}$$

and it follows by the minimization property of x_α^δ and $\alpha \ge 1$

$$\|x_\alpha^\delta - \bar{x}\|^2 \le \frac{1}{\alpha}\Phi_\alpha(x_\alpha^\delta) \le \frac{1}{\alpha}\Phi_\alpha(\bar{x}) \le \|y^\delta - F(\bar{x})\|^2 \ ,$$

and thus we have found for $\alpha \ge 1$

$$\|F'(x_\alpha^\delta)\| \le L\|y^\delta - F(\bar{x})\| + \|F'(\bar{x})\| \ . \quad (41)$$

Moreover, we have for $\alpha \le 1$

$$\|F'(x_\alpha^\delta)\| \le \|F'(x_\alpha^\delta) - F'(x_*)\| + \|F'(x_*) - F'(\bar{x})\| + \|F'(\bar{x})\| \ , \quad (42)$$

and the first term can be estimated as follows: By (7),

$$\|F'(x_\alpha^\delta) - F'(x_*)\| \le L\|x_\alpha^\delta - x_*\| \ .$$

For $\alpha \le 1$ we use estimate (12) and get

$$\begin{aligned} \|x_\alpha^\delta - x_*\| &\le \frac{\delta + \alpha\|\omega\|}{\sqrt{\alpha}\sqrt{1-L\|\omega\|}} \\ &\le \frac{\eta + \|\omega\|}{\sqrt{1-L\|\omega\|}}\sqrt{\alpha} \\ &\le \frac{2\varrho}{\sqrt{1-L\varrho}} \ . \end{aligned}$$

For the second term in (42) we obtain

$$\begin{aligned}\|F'(x_*) - F'(\bar{x})\| &\leq L\|\bar{x} - x_*\| = L\|F'(x_*)^*\omega\| \\ &\leq L\|\omega\|\|F'(\bar{x})\| + L\|\omega\|\|F'(x_*) - F'(\bar{x})\| \\ &\leq L\|\omega\|\|F'(\bar{x})\| + (L\|\omega\|)^2\|F(\bar{x})\| + (L\|\omega\|)^2\|\|F'(x_*) - F'(\bar{x})\| \\ &\vdots \\ &\leq \|F'(\bar{x})\| \sum_{j=1}^{n} (L\|\omega\|)^j + (L\|\omega\|)^n \|F'(x_*) - F'(\bar{x})\| \;.\end{aligned}$$

The right hand side converges because of $L\|\omega\| < 1$ for $n \to \infty$,

$$\|F'(x_*) - F'(\bar{x})\| \leq \|F'(\bar{x})\| \frac{1}{1 - L\|\omega\|} \leq \|F'(\bar{x})\| \frac{1}{1 - L\varrho} \;. \tag{43}$$

Thus we have found for $\alpha \leq 1$

$$\|F'(x_\alpha^\delta)\| \leq \frac{2L\varrho}{\sqrt{1 - L\varrho}} + \left(\frac{1}{1 - L\varrho} + 1\right) \|F'(\bar{x})\| \tag{44}$$

and combining (41), (44) yields (40).

□

To simplify the notation, we will define

$$a(t,h) := \langle F''(x_\alpha^\delta + th)(h,h), C_{x_\alpha^\delta}(t)(h,h)\rangle + 2\|\tilde{C}_{x_\alpha^\delta}(t)(h,h)\|^2 \;, \tag{45}$$

$$b(t,h) := \langle F'(x_\alpha^\delta)h, 4\tilde{C}_{x_\alpha^\delta}(t)(h,h) + F''(x_\alpha^\delta + th)(h,h)\rangle \;, \tag{46}$$

$$c(t,h) := 2\|F'(x_\alpha^\delta)h\|^2 + 2\alpha - \langle y^\delta - F(x_\alpha^\delta), F''(x_\alpha^\delta + th)(h,h)\rangle \;,$$

and have therefore

$$\varphi''_{\alpha,h}(t) = c(t,h) + tb(t,h) + t^2 a(t,h) \;. \tag{47}$$

Proposition 4.4 *Let $a(t,h)$, $b(t,h)$ and $c(t,h)$ be defined as in (45) and α be chosen such that $\delta = \eta\alpha$ with $\eta \leq \varrho$. Moreover, assume that γ is chosen such that*

$$\gamma + \frac{3}{2}L\|\omega\| \leq \gamma + \frac{3}{2}L\varrho < 1 \tag{48}$$

holds. Then $a(t,h)$, $b(t,h)$ and $c(t,h)$ can be estimated independently of t,h:

$$|a(t,h)| \leq \frac{3}{4}L^2 \tag{49}$$

$$|b(t,h)| \leq 3LK \tag{50}$$

$$c(t,h) > 2\gamma\alpha \;, \tag{51}$$

where K is defined in (42).

Proof:
According to (17), (27) and (37) we have with $\|h\| = 1$

$$
\begin{aligned}
\|C_{x_\alpha^\delta}(t)(h,h)\| &\leq \frac{1}{2}\int_0^1 (1-\tau)\|F''(x_\alpha^\delta + t\tau h)(h,h)\|\,d\tau \leq \frac{L}{4} \\
\|\tilde{C}_{x_\alpha^\delta}(t)(h,h)\| &\leq \frac{1}{2}\int_0^1 \|F''(x_\alpha^\delta + t\tau h)(h,h)\|\,d\tau \leq \frac{L}{2}
\end{aligned}
$$

and thus

$$
|a(t,h)| \leq \|F''(x_\alpha^\delta + th)(h,h)\|\|C_{x_\alpha^\delta}(t)(h,h)\| + 2\|\tilde{C}_{x_\alpha^\delta}(t)(h,h)\|^2 \leq \frac{3}{4}L^2 \ .
$$

Similarly,

$$
\begin{aligned}
|b(t,h)| &\leq \|F'(x_\alpha^\delta)\|\left(4\|\tilde{C}_{x_\alpha^\delta}(t)(h,h)\| + \|F''(x_\alpha^\delta + th)(h,h)\|\right) \leq 3L\|F'(x_\alpha^\delta)\| \\
&\overset{(42)}{\leq} 3LK \ .
\end{aligned}
$$

$c(t,h)$ can be estimated as follows:

$$
\begin{aligned}
c(t,h) - 2\gamma\alpha &= 2(\|F'(x_\alpha^\delta)h\|^2 + \alpha) - \langle y^\delta - F(x_\alpha^\delta), F''(x_\alpha^\delta + th)(h,h)\rangle - 2\gamma\alpha \\
&\geq 2\alpha - L\|y^\delta - F(x_\alpha^\delta)\| - 2\gamma\alpha \\
&\overset{(11)}{\geq} 2\alpha - L(\delta + 2\alpha\|\omega\|) - 2\gamma\alpha \\
&= 2\alpha - L(\eta\alpha + 2\alpha\|\omega\|) - 2\gamma\alpha \\
&\geq 2\left(1 - \frac{3}{2}L\varrho - \gamma\right)\alpha > 0 \ . \qquad (52)
\end{aligned}
$$

□

Theorem 4.5 *Let $\delta = \eta\alpha$, $\eta \leq \varrho$. The function $\varphi_{\alpha,h}(t)$, $\|h\| = 1$, is under assumption (48) a convex function for all*

$$
|t| \leq \frac{1}{L}\min\left\{\frac{1}{1+\sqrt{2}}\sqrt{\frac{8\kappa\alpha}{3}}\ , \frac{8\kappa\alpha}{3K(2+\sqrt{8})}\right\} =: r(\alpha) \ , \qquad (53)
$$

where κ and K are defined in (57), (42) . Especially holds

$$
\varphi''_{\alpha,h}(t) \geq 2\gamma\alpha \qquad (54)
$$

for $|t| \leq r(\alpha)$.

Proof:
It is

$$\varphi''_{\alpha,h}(t) = c(t,h) + tb(t,h) + t^2 a(t,h) \ .$$

We have to consider two cases:
1. Let $t > 0$. Then

$$\varphi''_{\alpha,h}(t) \geq c(t,h) - t|b(t,h)| - t^2|a(t,h)| \qquad \text{for } b(t,h) \leq 0 \ , \tag{55}$$

$$\varphi''_{\alpha,h}(t) \geq c(t,h) - t^2|a(t,h)| \qquad \text{for } b(t,h) > 0 \ . \tag{56}$$

2. Let $t \leq 0$. Then we have

$$\varphi''_{\alpha,h}(t) \geq c(t,h) - t^2|a(t,h)| \qquad \text{for } b(t,h) \leq 0 \ ,$$

$$\varphi''_{\alpha,h}(t) \geq c(t,h) - t|b(t,h)| - t^2|a(t,h)| \qquad \text{for } b(t,h) > 0 \ .$$

Thus it is sufficient to consider the first case only. From (56) follows with (49) and (52)

$$\varphi''_{\alpha,h}(t) - 2\gamma\alpha \geq 2\left(1 - \frac{3}{2}L\varrho - \gamma\right)\alpha - \frac{3}{4}L^2t^2 := p_1(t) \ .$$

Setting

$$\kappa = 1 - \frac{3}{2}L\varrho - \gamma \ , \tag{57}$$

then $p_1(t)$ has the zeros

$$t_{1,2} = \pm\sqrt{\frac{8\kappa}{3L^2}}\sqrt{\alpha} \ . \tag{58}$$

and because of $p_1(0) > 0$ holds

$$\varphi''_{\alpha,h}(t) \geq 2\gamma\alpha$$

for $|t| \leq |t_{1,2}|$. From (55) follows with (49), (50)

$$\varphi''_{\alpha,h}(t) - 2\gamma\alpha \geq 2\kappa\alpha - 3LKt - \frac{3}{4}L^2t^2 =: p_2(t) \ ,$$

and $p_2(t)$ has the zeros

$$t_{1,2} = \frac{1}{L}\left[-2K \pm \sqrt{4K^2 + \frac{8\kappa\alpha}{3}}\right] \ .$$

Now let $t_{min} = \min\{|t_1|, |t_2|\}$. Then

$$\begin{aligned} t_{min} &= \frac{1}{L}\left[-2K + \sqrt{4K^2 + \frac{8\kappa\alpha}{3}}\right] \\ &= \frac{8}{3L}\frac{\kappa\alpha}{2K + \sqrt{4K^2 + \frac{8}{3}\kappa\alpha}} \\ &\geq \frac{1}{L}\begin{cases} \frac{8\kappa\alpha}{3K(2+\sqrt{8})} & \text{if } \frac{8}{3}\kappa\alpha \leq 4K^2 \\ \frac{1}{1+\sqrt{2}}\sqrt{\frac{8}{3}\kappa\alpha} & \text{if } \frac{8}{3}\kappa\alpha > 4K^2 \ . \end{cases} \end{aligned} \tag{59}$$

Combining (58), (59) we have shown $\varphi''_{\alpha,h}(t) \geq 2\gamma\alpha$ for

$$\begin{aligned} |t| &\leq \frac{1}{L}\min\left\{\sqrt{\frac{8\kappa\alpha}{3}}, \frac{1}{1+\sqrt{2}}\sqrt{\frac{8\kappa\alpha}{3}}, \frac{8\kappa\alpha}{3K(2+\sqrt{8})}\right\} \\ &= \frac{1}{L}\min\left\{\frac{1}{1+\sqrt{2}}\sqrt{\frac{8\kappa\alpha}{3}}, \frac{8\kappa\alpha}{3K(2+\sqrt{8})}\right\} \end{aligned} \quad (60)$$

□

5 The steepest descent method

The steepest descent method is a widely used iteration method for minimizing a functional $\Phi : H \to \mathbb{R}$. Although it is sometimes slow in convergence, it seldom fails to converge to a minimum of the functional. For operators F which can be decomposed into $F(x) = Af + B(f, \mu)$, $x = (f, \mu)$, where A is a linear and B a bilinear operator, we have shown in [25] that steepest descent can be used to find a global minimizer as long as a starting value x_0, belonging to the area of convexity of the function $\varphi_{\alpha,h}(t)$, for the iteration exists. The techniques developed in [25] can in principle be used for a convergence proof of steepest descent method with twice continuous Fréchet–differentiable operators. However, there are some significant changes in the proofs. Additionally, we will need later on some quantitative estimates on convergence rates of the method and the determination of a step size parameter. Thus we will give in short a convergence analysis for the steepest descent method for minimizing the Tikhonov–functional $\Phi_\alpha(x)$.

The method is defined by

$$x_{k+1} = x_k + \beta_k \nabla\Phi_\alpha(x_k) \ , \quad (61)$$

where $\nabla\Phi_\alpha(x_k)$ denotes the direction of steepest descent of Φ_α at point x_k, and $\beta_k \in \mathbb{R}^+$ is a step size or scaling parameter. According to (29) we find

$$\nabla\Phi_\alpha(x_k) = 2(F'(x_k)^*(y^\delta - F(x_k)) - \alpha(x - \bar{x})) \ . \quad (62)$$

For the following, we might denote by

$$K_{r(\alpha)}(x_\alpha^\delta) := \left\{x \in X | x = x_\alpha^\delta + th, t \in \mathbb{R}, h \in X \text{ and } \|th\| \leq r(\alpha)\right\} \quad (63)$$

(for the definition of $r(\alpha)$, see (53)).

As in [25], we define

$$h_k := x_\alpha^\delta - x_k \ , \quad (64)$$

and functions $\varphi_1(t), \varphi_2(t)$ by

$$\varphi_1(t) = \Phi_\alpha(x_k + th_k) \tag{65}$$
$$\varphi_2(t) = \Phi_\alpha(x_\alpha^\delta - th_k) \ . \tag{66}$$

φ_1 and φ_2 can be rewritten as

$$\varphi_1(t) = \varphi_1(0) - 2\langle \nabla\Phi_\alpha(x_k), h_k\rangle t + c_1(t,h_k)\cdot t^2 + b_1(t,h_k)\cdot t^3 + a_1(t,h_k)\cdot t^4 \tag{67}$$
$$\varphi_2(t) = \varphi_2(0) + \underbrace{2\langle \nabla\Phi_\alpha(x_\alpha^\delta), h_k\rangle}_{=0} t + c_2(t,h_k)\cdot t^2 + b_2(t,h_k)\cdot t^3 + a_2(t,h_k)\cdot t^4 \tag{68}$$

where the coefficients $c_1(t,h_k), b_1(t,h_k), a_1(t,h_k), c_2(t,h_k), b_2(t,h_k), a_2(t,h_k)$ can be determined as in (29), e.g.

$$c_1(t,h_k) = \|F'(x_k)h_k\|^2 - 2\langle y^\delta - F(x_k), C_{x_k}(t)(h_k,h_k)\rangle + \alpha\|h_k\|^2 \tag{69}$$
$$b_1(t,h_k) = \langle F'(x_k)h_k, C_{x_k}(t)(h_k,h_k)\rangle \tag{70}$$
$$a_1(t,h_k) = \|C_{x_k}(t)(h_k,h_k)\|^2 \tag{71}$$

In contrary to the setting in [25] do these coefficients depend additionally on t, which is the main reason that we cannot completely use the analysis in [25]. The following two Propositions remain unchanged:

Proposition 5.1 *Let the assumptions of Theorem 4.5 hold, and let $\varphi_1(t)$ be defined as in (65) and $x_k \in K_{r(\alpha)}(x_\alpha^\delta)$. Then we have*

$$\varphi_1'(t) < 0 \ for\ t \in [0,1)\ ,$$
$$\varphi_1'(t) = 0 \ for\ t = 0\ ,$$

and

$$\varphi_1''(t) \geq 2\gamma\alpha\|h_k\|^2\ ,\ for\ t\in[0,1]\ . \tag{72}$$

Proposition 5.2 *Assume that $x_k \in K_{r(\alpha)}(x_\alpha^\delta)$. Then an interval $I = (0,\beta_0]$,*

$$\beta_0 = \frac{\langle \nabla\Phi_\alpha(x_k), h_k\rangle}{\|\nabla\Phi_\alpha(x_k)\|^2}\ , \tag{73}$$

exists such that the iterate $x_{k+1} = x_k + \beta_k\nabla\Phi_\alpha(x_k)$ is closer to x_α^δ than x_k for $\beta_k \in I$:

$$\|x_\alpha^\delta - x_{k+1}\| < \|x_\alpha^\delta - x_k\|\ .$$

Especially, $x_{k+1} \in K_{r(\alpha)}(x_\alpha^\delta)$.

For a proof, we refer to Propositions 2.4. and 2.5. in [25].

Proposition 5.3 *Let the assumptions of Theorem 4.5 hold, x_j, $j = 0, ..., k+1$ be the steepest descent iterates for minimizing Φ_α and h_k be defined as in (64). Moreover, assume $x_j \in K_{r(\alpha)}(x_\alpha^\delta))$ for $j = 0, ..., k$. Then*

$$\langle \nabla\Phi_\alpha(x_k), h_k\rangle \geq \gamma\alpha\|h_k\|^2\ . \tag{74}$$

Proof:
Let $\varphi_1(t)$, $\varphi_2(t)$ be defined as in (65)–(68) and $t \in [0,1]$, and $g_k(t)$ as $g(t)$ in (30) with x replaced by x_k and h replaced by h_k. Then $\varphi_1(t)$ is computed by (cf. (39))

$$\begin{aligned}\varphi_1(t) &= \varphi_1(0) - 2t\langle \nabla\Phi_\alpha(x_k), h_k\rangle + t^2(\|F'(x_k)h_k\|^2 + \alpha\|h_k\|^2) \\ &\quad -2\langle y^\delta - F(x_k), g_k(t)\rangle + 2\langle tF'(x_k)h_k, g_k(t)\rangle + \|g_k(t)\|^2 ,\end{aligned}$$

and we get by Taylor's formula

$$\begin{aligned}\varphi_1'(t) &= \varphi_1'(0) + t\int_0^1 \varphi_1''(\theta t)\, d\theta \\ &= -2\langle \nabla\Phi_\alpha(x_k), h_k\rangle + t\int_0^1 \varphi_1''(\theta t)\, d\theta .\end{aligned}$$

Now we have $\theta t \in [0,1]$ for $t \in [0,1]$, $\varphi_1'(t) \le 0$ and $\varphi_1''(\theta t) \ge 2\gamma\alpha\|h_k\|^2$ for these t (cf. (72)). Thus we conclude

$$2\langle \nabla\Phi_\alpha(x_k), h_k\rangle = -\varphi_1'(t) + t\int_0^1 \varphi_1''(\theta t)\, d\theta \ge 2t\gamma\alpha\|h_k\|^2$$

for all $t \in [0,1]$, and esp. $t = 1$ yields (74).

□

Proposition 5.4 *Let the assumptions of Theorem 4.5 hold, and let x_j, $j = 0, ..., k+1$ be the steepest descent iterates for minimizing Φ_α. If $x_j \in K_{r(\alpha)}(x_\alpha^\delta)$ for $j = 0, ..., k$ and the scaling parameter β_k is chosen such that*

$$\beta_k \le \min\left\{ \frac{\gamma\alpha}{\|\nabla\Phi_\alpha(x_k)\|^2}, \frac{16\gamma\alpha}{16K^2 + 16\alpha + 24\varrho\alpha L + 8KL + L^2} \frac{(\Phi_\alpha(x_k) - \phi_{min,k})}{\|\nabla\Phi_\alpha(x_k)\|^2} \right\} \tag{75}$$

holds, then the new iterate x_{k+1} is closer to x_α^δ than x_k. Here,

$$\phi_{min,k} = \min\{\Phi_\alpha(x_k) + t\nabla\Phi_\alpha(x_k)) :\ t \in \mathbb{R}^+\} . \tag{76}$$

Especially, $x_{k+1} \in K_{r(\alpha)}(x_\alpha^\delta)$.

Proof:
Let φ_1, φ_2 be defined as in (65)–(68). We will first estimate $\|h_k\|$ from below by $\varphi_1(0) - \varphi_1(1)$. We have $\varphi_1(t) = \varphi_2(1-t)$. Keeping in mind that $\nabla\Phi_\alpha(x_\alpha^\delta) = 0$ holds, we get from (68)

$$\begin{aligned}\varphi_1(0) - \varphi_1(1) &= \varphi_2(1) - \varphi_2(0) \\ &= c_2(t,h_k) + b_2(t,h_k) + a_2(t,h_k) .\end{aligned} \tag{77}$$

By setting

$$C_{k,x_\alpha^\delta}(t)(h_k,h_k) := \frac{1}{2}\int_0^1 (1-\tau)F''(x_\alpha^\delta + t\tau h_k)(h_k,h_k)\,d\tau \ , \tag{78}$$

the coefficients in (77) can be determined by (cf. (29))

$$c_2(t,h_k) = \|F'(x_\alpha^\delta)h_k\|^2 + \alpha\|h_k\|^2 - 2\langle y^\delta - F(x_\alpha^\delta), C_{k,x_\alpha^\delta}(t)(h_k,h_k)\rangle \tag{79}$$

$$b_2(t,h_k) = 2\langle F'(x_\alpha^\delta)h_k, C_{k,x_\alpha^\delta}(t)(h_k,h_k)\rangle \tag{80}$$

$$a_2(t,h_k) = \|C_{k,x_\alpha^\delta}(t)(h_k,h_k)\|^2 \ . \tag{81}$$

Using the estimate (40), $\delta = \eta\varrho$, $\eta \le \varrho$ and

$$\|C_{k,x_\alpha^\delta}(t)(h_k,h_k)\| \overset{(17)}{\le} \frac{L}{2}\|\int_0^1 (1-\tau)\,d\tau\|h_k\|^2 = \frac{L}{4}\|h_k\|^2 \tag{82}$$

$$\|y^\delta - F(x_\alpha^\delta)\| \overset{(11)}{\le} \delta + 2\alpha\|\omega\| \le \eta\alpha + 2\alpha\|\omega\| \le 3\varrho\alpha \ , \tag{83}$$

we obtain

$$|c_2(t,h_k)| \le (K^2 + \alpha + \frac{3}{2}\varrho\alpha L)\|h_k\|^2 =: \bar{c}_2\|h_k\|^2 \tag{84}$$

$$|b_2(t,h_k)| \le \frac{KL}{2}\|h_k\|^3 =: \bar{b}_2\|h_k\|^3 \tag{85}$$

$$|a_2(t,h_k)| \le \frac{L^2}{16}|\|h_k\|^4 =: \bar{a}_2\|h_k\|^4 \ . \tag{86}$$

Altogether this yields

$$\varphi_1(0) - \varphi_1(1) \le \bar{c}_2\|h_k\|^2 + \bar{b}_2\|h_k\|^3 + \bar{a}_2\|h_k\|^4 \ .$$

The minimal value of Φ_α, $\Phi_\alpha(x_\alpha^\delta) = \varphi_1(1)$, is usually not known, but $\phi_{min,k}$, defined in (76), is computable. In case of $\|h_k\| \le 1$ we find

$$\begin{aligned} \varphi_1(0) - \phi_{min,k} &\le \varphi_1(0) - \varphi_1(1) \\ &\le (K^2 + \alpha + \frac{3}{2}\varrho\alpha L)\|h_k\|^2 + \frac{KL}{2}\|h_k\|^3 + \frac{L^2}{16}\|h_k\|^4 \\ &\le (K^2 + \alpha + \frac{3}{2}\varrho\alpha L + \frac{KL}{2} + \frac{L^2}{16})\|h_k\|^2 \ . \end{aligned} \tag{87}$$

Thus it follows by using (74) for $\|h_k\| \le 1$

$$\langle \nabla\Phi_\alpha(x_k), h_k\rangle \ge \frac{16\gamma\alpha}{16K^2 + 16\alpha + 24\varrho\alpha L + 8KL + L^2}(\varphi_1(0) - \phi_{min,k}) \ .$$

In case of $\|h_k\| > 1$ we get from (74)

$$\langle \nabla\Phi_\alpha, h_k\rangle \ge \gamma\alpha \ .$$

Inserting the above estimates for $\langle \nabla\Phi_\alpha(x_k), h_k\rangle$ in (73) shows that $\beta_k \le \beta_0$, and, by Proposition 5.2, the new iterate is therefore closer to x_α^δ than the old one.

□

Proposition 5.5 *Let the assumptions of Theorem 4.5 hold, $K_{r(\alpha)}(x_\alpha^\delta)$ be defined as in (63), $x_0 \in K_{r(\alpha)}(x_\alpha^\delta)$, and $\{x_k\}_{k\in\mathbb{N}}$ be the iterates of steepest descent for the Tikhonov–functional with β_k chosen according to (75). Then, there exists a constant M, defined in (93), such that the second derivative of*

$$\phi_k(t) := \Phi_\alpha\left(x_k + t\nabla\Phi_\alpha(x_k)\right) \qquad t \in [0,1]\ ,$$

is bounded by

$$|\phi_k''(t)| \le M\|\nabla\Phi_\alpha(x_k)\|^2\ .$$

Proof:
According to the choice of β_k all iterates are in $K_{r(\alpha)}(x_\alpha^\delta)$. The definition (62) of $\nabla\Phi_\alpha(x_k)$ shows that $\|\nabla\Phi_\alpha(x_k)\|$ bounded in $K_{r(\alpha)}(x_\alpha^\delta)$ by a constant κ: It is

$$\|\nabla\Phi_\alpha(x_k)\| \le \|F'(x_k)\|\|y^\delta - F(x_k)\| + \alpha\|x_k - \bar{x}\|\ ,$$

and

$$\begin{aligned}
\|F'(x_k)\| &\le \|F'(x_k) - F'(x_\alpha^\delta)\| + \|F'(x_\alpha^\delta)\| \overset{(7),(40)}{\le} L\|x_\alpha^\delta - x_k\| + K \\
&\le Lr(\alpha) + K =: \kappa_1 && (88)\\
\|y^\delta - F(x_k)\| &\overset{(83)}{\le} 3\varrho\alpha + \|F(x_\alpha^\delta) - F(x_k)\| \overset{(16)-(17)}{\le} 3\varrho\alpha + \|F'(x_\alpha^\delta)\|\|x_\alpha^\delta - x_k\| + \frac{L}{4}\|x_\alpha^\delta - x_k\|^2 \\
&\overset{(40)}{\le} 3\delta + Kr(\alpha) + \frac{L}{4}r(\alpha)^2 =: \kappa_2 && (89)\\
\alpha^2\|x_k - \bar{x}\|^2 &\le \alpha^2(\|x_k - x_\alpha^\delta\|^2 + \|x_\alpha^\delta - \bar{x}\|^2 + 2\|x_k - x_\alpha^\delta\|\|x_\alpha^\delta - \bar{x}\|) \\
&\le \alpha^2 r(\alpha)^2 + \alpha\Phi_\alpha(x_\alpha^\delta) + 2\alpha^2\max\{\|x_k - x_\alpha^\delta\|^2, \|x_\alpha^\delta - \bar{x}\|^2\} \\
&\le \alpha^2 r(\alpha)^2 + \alpha\Phi_\alpha(\bar{x}) + 2\alpha\max\{\alpha r(\alpha)^2, \Phi_\alpha(\bar{x})\} \\
&\le \alpha^2 r(\alpha)^2 + \alpha\|y^\delta - F(\bar{x})\| + 2\alpha\max\{\alpha r(\alpha), \|y^\delta - F(\bar{x})\|\} =: \kappa_3^2 && (90)
\end{aligned}$$

Thus

$$\|\nabla\Phi_\alpha(x_k)\| \le \kappa := \kappa_1\cdot\kappa_2 + \kappa_3 \tag{91}$$

holds for all k. Defining $C_{k,x_\alpha^\delta}(t)(\nabla\Phi_\alpha(x_k), \nabla\Phi_\alpha(x_k))$ as in (78) with h_k replaced by $\nabla\Phi_\alpha(x_k)$ and

$$\tilde{C}_{k,x_\alpha^\delta}(t)(\nabla\Phi_\alpha(x_k), \nabla\Phi_\alpha(x_k)) := \frac{1}{2}\int_0^1 F''(x_k + t\tau\nabla\Phi_\alpha(x_k))(\nabla\Phi_\alpha(x_k), \nabla\Phi_\alpha(x_k))\, d\tau\ , \tag{92}$$

we find (cf. (38))

$$\begin{aligned}\phi_k''(t) &= 2\|F'(x_k)\nabla\Phi_\alpha(x_k)\|^2 - \langle y^\delta - F(x_k), F''(x_k + t\nabla\Phi_\alpha(x_k))(\nabla\Phi_\alpha(x_k), \nabla\Phi_\alpha(x_k))\rangle \\ &\quad +2\alpha\|\nabla\Phi_\alpha(x_k)\|^2 + 4t\langle F'(x_k)\nabla\Phi_\alpha(x_k), \tilde{C}_{k,x_\alpha^\delta}(t)(\nabla\Phi_\alpha(x_k), \nabla\Phi_\alpha(x_k)) \\ &\quad +t\langle F'(x_k)\nabla\Phi_\alpha(x_k), F''(x_k + t\nabla\Phi_\alpha(x_k))(\nabla\Phi_\alpha(x_k), \nabla\Phi_\alpha(x_k))\rangle \\ &\quad +t^2\langle F''(x_k + t\nabla\Phi_\alpha(x_k))(\nabla\Phi_\alpha(x_k), \nabla\Phi_\alpha(x_k)), C_{k,x_\alpha^\delta}(t)(\nabla\Phi_\alpha(x_k), \nabla\Phi_\alpha(x_k))\rangle \\ &\quad +2t^2\|\tilde{C}_{k,x_\alpha^\delta}(t)(\nabla\Phi_\alpha(x_k), \nabla\Phi_\alpha(x_k))\|^2 .\end{aligned}$$

The norm of $C_{k,x_\alpha^\delta}(t)(\nabla\Phi_\alpha(x_k), \nabla\Phi_\alpha(x_k))$, $\tilde{C}_{k,x_\alpha^\delta}(t)(\nabla\Phi_\alpha(x_k), \nabla\Phi_\alpha(x_k))$ can be estimated similarly as in (82),

$$\begin{aligned}\|C_{k,x_\alpha^\delta}(t)(\nabla\Phi_\alpha(x_k), \nabla\Phi_\alpha(x_k))\| &\le \frac{L}{4}\|\nabla\Phi_\alpha(x_k)\|^2 , \\ \|\tilde{C}_{k,x_\alpha^\delta}(t)(\nabla\Phi_\alpha(x_k), \nabla\Phi_\alpha(x_k))\| &\le \frac{L}{2}\|\nabla\Phi_\alpha(x_k)\|^2.\end{aligned}$$

Using the above given estimates and (17), (88)-(91) we can finally estimate $|\phi_k''(t)|$ for $t \in [0,1]$ by

$$\begin{aligned}|\phi_k''(t)| &\le (\kappa_1^2 + L\kappa_2 + 2\alpha)\|\nabla\Phi_\alpha(x_k)\|^2 + 3t\kappa_1 L\|\nabla\Phi_\alpha(x_k)\|^3 \\ &\quad +t^2\frac{3L^2}{4}\|\nabla\Phi_\alpha(x_k)\|^4 \\ &\le \left(\kappa_1^2 + L\kappa_2 + 2\alpha + 3\kappa_1\kappa + \frac{3L^2}{4}\kappa^2\right)\|\nabla\Phi_\alpha(x_k)\|^2 \\ &=: M\|\nabla\Phi_\alpha(x_k)\|^2 . \end{aligned} \tag{93}$$

□

Now all ingredients for the final convergence proof of the steepest descent method have been collected:

Theorem 5.6 *Let the assumptions of Theorem 4.5 hold,* $K_{r(\alpha)}(x_\alpha^\delta)$ *be defined as in (63),* $x_0 \in K_{r(\alpha)}(x_\alpha^\delta)$*, and* $\{x_k\}_{k\in\mathbb{N}}$ *be the iterates of steepest descent for the Tikhonov functional* Φ_α *with*

$$\beta_k = \min\left\{\frac{\gamma\alpha}{\|\nabla\Phi_\alpha(x_k)\|^2}, \frac{16\gamma\alpha}{16K^2 + 16\alpha + 24\varrho\alpha L + 8KL + L^2}\frac{(\Phi_\alpha(x_k) - \phi_{min,k})}{\|\nabla\Phi_\alpha(x_k)\|^2}, \frac{1}{M}, 1\right\} , \tag{94}$$

where $\phi_{min,k}$ *and* M *are defined as in (76), (93). Then* x_k *converges to a global minimizer of* Φ_α*:*

$$x_k \to x_\alpha^\delta \quad \textit{for } k \to \infty .$$

The proof is similar to the proof of Theorem 2.9 in [25].
At the end of this section we might remark that (74) provides a control over the accuracy of the k-th iterate. Indeed, we conclude

$$\|x_\alpha^\delta - x_k\| \leq \frac{1}{\gamma\alpha}\|\nabla\Phi_\alpha(x_k)\| \ , \tag{95}$$

which will be a useful estimate for the algorithm to be defined later on.

6 Continuity of the map $\alpha \to x_\alpha^\delta$ and Morozov's Discrepancy principle

The choice of the regularization parameter is of highest interest in the theory of regularization methods. A posteriori parameter choice strategies have been used successfully for linear inverse problems. Well known is Morozov's discrepancy principle [18], where the regularization parameter is chosen such that

$$\|y^\delta - F(x_\alpha^\delta)\| = c\delta \tag{96}$$

holds. For general nonlinear operators, it is a well known fact for Tikhonov–regularization that the mapping $\alpha \to x_\alpha^\delta$ as well as the mapping $\alpha \to \|y^\delta - F(x_\alpha^\delta)\|$ might be discontinuous [30, 4], and thus a regularization parameter α with (96) might not exist. It is therefore of interest to give conditions under which the continuity of the above mentioned mappings can be guaranteed. Kravaris and Seinfeld [4] and Scherzer [27] gave conditions, under which a regularization parameter with (96) exist. However, for practical applications it is sometimes difficult to prove the required conditions, and thus we introduced in [24] a modified discrepancy principle, and showed that this principle can be used under relatively mild conditions. Here, we will focus on the continuity of the map $\alpha \to x_\alpha^\delta$. We will show that the map is continuous for twice Fréchet differentiable operators if only a smoothness condition $x_* - \bar{x} = F'(x_*)^*\omega$ with $\|\omega\|$ small enough holds. This actually gives a new focus on the smoothness condition for nonlinear problems: For linear problems, this condition was needed to obtain convergence convergence rates. All the other above discussed properties are obtained automatically by the linearity of the problem. For nonlinear problems, we will get additionally the continuity of the mapping $\alpha \to x_\alpha^\delta$, the existence of a regularization parameter with (96) and the uniqueness of the minimizer x_α^δ of the Tikhonov–functional.

If the solution x_* of the equation $F(x) = y$ fulfills a smoothness condition, then we will show for a sequence $\alpha_k \to \alpha$ for $k \to \infty$ that the sequence $x_{\alpha_k}^\delta$ of minimizers of the Tikhonov–functional Φ_{α_k} converges towards x_α^δ:

Theorem 6.1 *Let F be a twice Fréchet differentiable operator with (7). Moreover, assume that the solution x_* of $F(x) = y$ fulfills the smoothness condition*

$$x_* - \bar{x} = F'(x_*)^*\omega \tag{97}$$

with

$$\|\omega\| \leq \varrho < \frac{1}{L} \ , \tag{98}$$

and that a sequence $\{\alpha_k\}$ *with* $\lim_{k\to\infty}\alpha_k = \alpha = \delta/\eta$, $\alpha_k > 0$, $\eta \le \varrho$ *is given. If*

$$L\left[\eta + 2\varrho + \frac{\eta+\varrho}{4\sqrt{1-\varrho L}}\right] < 1 \tag{99}$$

holds, then the estimate

$$\|x^\delta_\alpha - x^\delta_{\alpha_k}\| \le \frac{|\alpha-\alpha_k|(\eta+2\varrho)}{2(1-p_k)\sqrt{q_k\alpha}} \ . \tag{100}$$

with p_k, q_k *given in (113), (110), yields for* k *large enough. Especially converges* $x^\delta_{\alpha_k}$ *to* x^δ_α *for* $k\to\infty$.

Proof:
For the following, we set

$$\begin{aligned} A &:= F'(x^\delta_\alpha) \\ A_k &:= F'(x^\delta_{\alpha_k}) \ , \end{aligned}$$

and get with (18)

$$A_k = A + \int_0^1 F''(x^\delta_\alpha + \tau(x^\delta_{\alpha_k} - x^\delta_\alpha))(x^\delta_{\alpha_k} - x^\delta_\alpha,\ \cdot\ )\,d\tau \ .$$

As in (20) we define the linear operator B_k by

$$B_k(\cdot) = \int_0^1 F''(x^\delta_\alpha + \tau(x^\delta_{\alpha_k} - x^\delta_\alpha))(x^\delta_{\alpha_k} - x^\delta_\alpha,\ \cdot\ )\,d\tau \ , \tag{101}$$

with

$$\|B_k\| \overset{(17)}{\le} L\|x^\delta_{\alpha_k} - x^\delta_\alpha\| \ , \tag{102}$$

and find therefore

$$A_k^* = A^* + B_k^* \ . \tag{103}$$

Moreover, let us define

$$C(h,h) := \int_0^1 \frac{(1-\tau)}{2} F''(x^\delta_\alpha + \tau h)(h,h)\,d\tau \ . \tag{104}$$

We conclude from (17)

$$\|C(h,h)\| \le \frac{L}{4}\|h\|^2 \ . \tag{105}$$

The necessary condition for a minimum of the Tikhonov–functional $\Phi_\alpha(x)$ reads

$$\alpha(x-\bar{x}) = F'(x)^*(y^\delta - F(x)) \ .$$

Thus we have with (103)

$$\begin{aligned}
x^\delta_{\alpha_k} - x^\delta_\alpha &= x^\delta_{\alpha_k} - \bar{x} + \bar{x} - x^\delta_\alpha \\
&= \frac{1}{\alpha_k} A_k^*(y^\delta - F(x^\delta_{\alpha_k})) - \frac{1}{\alpha} A^*(y^\delta - F(x^\delta_{\alpha_k})) \\
&= \frac{1}{\alpha_k}(A + B_k)^*(y^\delta - F(x^\delta_{\alpha_k})) - \frac{1}{\alpha} A^*(y^\delta - F(x^\delta_\alpha)) \\
&= \frac{\alpha - \alpha_k}{\alpha\alpha_k} A^* y^\delta + A^* \left(\frac{1}{\alpha} F(x^\delta_\alpha) - \frac{1}{\alpha_k} F(x^\delta_{\alpha_k}) \right) + \frac{1}{\alpha_k} B_k^*(y^\delta - F(x^\delta_{\alpha_k})) \\
&\overset{(16)}{=} \frac{\alpha - \alpha_k}{\alpha\alpha_k} A^*(y^\delta - F(x^\delta_\alpha)) + \frac{1}{\alpha_k} B_k^*(y^\delta - F(x^\delta_{\alpha_k})) \\
&\quad - \frac{1}{\alpha_k} A^* \left(A(x^\delta_{\alpha_k} - x^\delta_\alpha) + C(x^\delta_{\alpha_k} - x^\delta_\alpha, x^\delta_{\alpha_k} - x^\delta_\alpha) \right) ,
\end{aligned}$$

and we have shown

$$(\alpha_k I + A^*A)(x^\delta_{\alpha_k} - x^\delta_\alpha) = \frac{\alpha - \alpha_k}{\alpha} A^*(y^\delta - F(x^\delta_\alpha)) + B_k^*(y^\delta - F(x^\delta_{\alpha_k})) - A^*C(x^\delta_{\alpha_k} - x^\delta_\alpha, x^\delta_{\alpha_k} - x^\delta_\alpha) . \tag{106}$$

Now, by using the estimates

$$\begin{aligned}
\|(\alpha_k I + A^*A)^{-1} A^*\| &\leq \frac{1}{2\sqrt{\alpha_k}} \\
\|(\alpha_k I + A^*A)^{-1}\| &\leq \frac{1}{\alpha_k}
\end{aligned}$$

and (105), (102) we obtain

$$\begin{aligned}
\|x^\delta_{\alpha_k} - x^\delta_\alpha\| &\leq \frac{|\alpha - \alpha_k|}{2\alpha\sqrt{\alpha_k}} \|y^\delta - F(x^\delta_\alpha)\| + \frac{1}{\alpha_k} \|B_k\| \|y^\delta - F(x^\delta_\alpha)\| + \frac{1}{2\sqrt{\alpha_k}} \|C(x^\delta_{\alpha_k} - x^\delta_\alpha, x^\delta_{\alpha_k} - x^\delta_\alpha)\| \\
&\leq \frac{|\alpha - \alpha_k|}{2\alpha\sqrt{\alpha_k}} \|y^\delta - F(x^\delta_\alpha)\| + \frac{L}{\alpha_k} \|y^\delta - F(x^\delta_{\alpha_k})\| \|x^\delta_{\alpha_k} - x^\delta_\alpha\| + \frac{L}{8\sqrt{\alpha_k}} \|x^\delta_{\alpha_k} - x^\delta_\alpha\|^2
\end{aligned} \tag{107}$$

Setting $\delta = \eta\alpha$, (11), (12) yields

$$\|y^\delta - F(x^\delta_\alpha)\| \leq (\eta + 2\varrho)\alpha \tag{108}$$

$$\|x^\delta_\alpha - x_*\| \leq \frac{\eta + \varrho}{\sqrt{1 - \varrho L}} \sqrt{\alpha} , \tag{109}$$

and, by setting

$$q_k := \frac{\alpha_k}{\alpha} \tag{110}$$

we get

$$\|y^\delta - F(x^\delta_{\alpha_k})\| \leq (\eta + 2q_k\varrho)\alpha \tag{111}$$

$$\begin{aligned}
\|x^\delta_{\alpha_k} - x_*\| &\leq \frac{\eta\alpha + \alpha_k\varrho}{\sqrt{\alpha_k}\sqrt{1 - \varrho L}} \\
&\leq \frac{\eta + q_k\varrho}{\sqrt{q_k}\sqrt{1 - \varrho L}} \sqrt{\alpha} .
\end{aligned} \tag{112}$$

Inserting (108)-(112) in (107) yields

$$\begin{aligned}\|x^{\delta}_{\alpha_k} - x^{\delta}_{\alpha}\| &\leq \frac{|\alpha-\alpha_k|}{2\alpha\sqrt{\alpha_k}}(\eta+2\varrho)\alpha + \frac{L\sqrt{\alpha}}{8\sqrt{\alpha_k}}\left(\frac{\eta+\varrho}{\sqrt{1-\varrho L}} + \frac{\eta+q_k\varrho}{\sqrt{q_k}\sqrt{1-\varrho L}}\right)\|x^{\delta}_{\alpha_k} - x^{\delta}_{\alpha}\| \\ &\quad + \frac{L(\eta+2q_k\varrho)\alpha}{\alpha_k}\|x^{\delta}_{\alpha_k} - x^{\delta}_{\alpha}\| \\ &\leq \frac{|\alpha-\alpha_k|(\eta+2\varrho)}{2\sqrt{q_k\alpha}} + \frac{L}{8q_k}\left(\frac{\sqrt{q_k}(\eta+\varrho)+\eta+q_k\varrho}{\sqrt{1-\varrho L}}\right)\|x^{\delta}_{\alpha_k} - x^{\delta}_{\alpha}\| \\ &\quad + \frac{L(\eta+2q_k\varrho)}{q_k}\|x^{\delta}_{\alpha_k} - x^{\delta}_{\alpha}\| \ .\end{aligned}$$

By setting

$$p_k := \frac{L}{q_k}\left[\eta + 2q_k\varrho + \frac{\sqrt{q_k}(\eta+\varrho)+\eta+q_k\varrho}{8\sqrt{1-\varrho L}}\right] \tag{113}$$

we find

$$(1-p_k)\|x^{\delta}_{\alpha_k} - x^{\delta}_{\alpha}\| \leq \frac{|\alpha-\alpha_k|(\eta+2\varrho)}{2\sqrt{q_k\alpha}} \ .$$

Because of $q_k \to 1$ for $k \to \infty$ holds with (99)

$$\lim_{k\to\infty} p_k = L\left[\eta + 2\varrho + \frac{\eta+\varrho}{4\sqrt{1-\varrho L}}\right] < 1 \ .$$

Consequently, there exists $\bar{p}$ with $0 < \bar{p} \leq (1-p_k)$ for k large enough and (100) holds.

□

Conclusion 6.2 *The Tikhonov–functional Φ_α has a unique minimizer under the conditions of Theorem 6.1. Moreover, let the constant c in Morozov's discrepancy principle chosen such that $c \geq 3$ holds, and assume*

$$\|y^\delta - F(\bar{x})\| > c\delta \ . \tag{114}$$

Then there exists a regularization parameter α such that (96) yields.

Proof:
In Theorem 6.1, x^{δ}_{α} was denoting an arbitrary minimizer of the Tikhonov–functional. Thus, a sequence $\{x^{\delta}_{\alpha_k}\}$ of minimizers converges to every minimizer of Φ_α, which means the minimizer has to be unique.
In [24] was shown $x^{\delta}_{\alpha} \to \bar{x}$ for $\alpha \to \infty$, and because of (114) there exists α_0 with $\|y^\delta - F(x^{\delta}_{\alpha_0})\| \geq c\delta$. Setting $\alpha_{min} = \delta/\varrho$, $\alpha = \delta/\eta$ with $\eta \leq \varrho$ we have $\alpha_{min} \leq \alpha$. From (11) follows

$$\|y^\delta - F(x^{\delta}_{\alpha_{min}})\| \leq 3\delta \leq c\delta \ .$$

The function $\psi(\alpha) = \|y^\delta - F(x^{\delta}_{\alpha})\|$ is monotone increasing and, according to Theorem 6.1, continuous for $\alpha \in [\alpha_{min}, \alpha_0]$, and thus there exists a regularization parameter with (96).

□

Condition (99) means that η and ϱ have to be small enough. Especially, because of $\delta = \eta\alpha$, α should not be chosen to small compared to δ. This reflects the well known fact that a too small α usually leads to a bad reconstruction quality. We might give a simple criterion for (99).

Proposition 6.3 *For $\eta \leq \varrho$, condition (99) is fulfilled whenever*

$$L\varrho \leq 0.278 \tag{115}$$

holds.

Proof:
For $\eta \leq \varrho$ we have

$$L\left[\eta + 2\varrho + \frac{\eta+\varrho}{4\sqrt{1-\varrho L}}\right] \leq 3L\varrho + \frac{L\varrho}{2\sqrt{1-\varrho L}} \ .$$

Setting $x = L\varrho$, the right hand side is smaller than 1 if

$$f(x) = 3x + \frac{x}{2\sqrt{1-x}} - 1 < 0$$

holds. Finding the zeros of $f(x)$ is equivalent to finding the roots of a cubic polynomial, and we get $f(x) \leq 0$ if $x \leq 0.278$.

□

7 The TIGRA–algorithm

As earlier mentioned, the TIGRA–algorithm will be a combination of the steepest descent method for the minimization of the Tikhonov–functional and an optimization routine for finding a regularization parameter α such that Morozov's discrepancy principle holds. The algorithm is defined as follows:

- Given y^δ with $\|y^\delta - y\| \le \delta$
- Choose $\alpha_0,\ q < 1$ and $x_0 \in K_r(x^\delta_{\alpha_0}),$ with

$$\|y^\delta - F(x^\delta_{\alpha_0})\| > 5\delta$$

- $k = -1,\ l = 0,\ x^\delta_{-1,0} = x_0$
- **Repeat** (116)

$$k = k + 1$$

$$x^\delta_{k,0} = x^\delta_{k-1,l} \tag{117}$$

$$\alpha_k = q^k \alpha_0 \tag{118}$$

$$l = 0,$$

While $\|\nabla \Phi_{\alpha_k}(x^\delta_{k,l})\| > \ TOL(k)$ (119)

$$x^\delta_{k,l+1} = x^\delta_{k,l} + \beta_{k,l} \nabla \Phi_{\alpha_k}(x^\delta_{k,l}) \tag{120}$$

$$l = l + 1$$

end (121)

until $\|y^\delta - F(x^\delta_{k,l})\| \le 5\delta$ (122)

The algorithm consists of an inner iteration (119)–(121) and an outer iteration (116)–(122). The main idea is that we choose α_0 such that $x_0 \in K_{r(\alpha_0)}(x^\delta_{\alpha_0})$ holds, and the inner iteration will converge towards $x^\delta_{\alpha_0}$. If q is chosen properly, then the last inner iterate belongs to $K_{r(\alpha_1)}(x^\delta_{\alpha_1})$, and the inner iteration for $k = 1$ might converge to $x^\delta_{\alpha_1}$ and so forth. Subsequently, we compute approximations for the Tikhonov–minimizers $x^\delta_{\alpha_k}$, and will show that the last iterate is a reasonable approximation to the solution x_*. In the following, we will investigate how the parameters α_0, q and TOL have to be determined such that the inner as well as the outer iteration terminates and that the final iterate is s good approximation to the solution. We might assume that a starting value x_0 might be known with

$$\|x_0 - x_*\| \le \lambda\ . \tag{123}$$

In general, one would use the best known approximation x_0 as an a priori guess for the Tikhonov–functional as well, i.e. one would set $\bar{x} = x_0$. But the analysis of the algorithm will be the same if we distinguish between x_0 and $\bar{x}$. First we will show that there exist a regularization parameter α such that $x_0 \in K_{r(\alpha)}(x^\delta_\alpha)$ holds.

Proposition 7.1 *Assume that (115) holds, and that x_0 with (123) is given. Moreover, assume that the parameter γ (cf (48)) fulfills*

$$\gamma = \left(1 - \frac{3}{2} L\varrho\right) \tau\ , \tau \le 0.5\ . \tag{124}$$

If α is chosen such that (126) holds, then

$$x_0 \in K_{r(\alpha)}(x_\alpha^\delta) \ . \tag{125}$$

<u>Proof:</u>
Let $\alpha > 0$ be given. From (12) and (123) we get

$$\begin{aligned}\|x_\alpha^\delta - x_0\| &\le \|x_\alpha^\delta - x_*\| + \|x_* - x_0\| \\ &\le \frac{\delta + \alpha\|\omega\|}{\sqrt{\alpha}\sqrt{1 - L\|\omega\|}} + \lambda \\ &\le \frac{\delta + \alpha\varrho}{\sqrt{\alpha}\sqrt{1 - L\varrho}} + \lambda\end{aligned}$$

According to (53), $x_0 \in K_{r(\alpha)}(x_\alpha^\delta)$ holds for α large enough if

$$\|x_0 - x_\alpha^\delta\| \le \frac{1}{L(1+\sqrt{2})}\sqrt{\frac{8\kappa\alpha}{3}}$$

(cf. (53)) , which is true if

$$\lambda + \frac{\delta + \alpha\varrho}{\sqrt{\alpha}\sqrt{1 - L\varrho}} \le \frac{1}{L(1+\sqrt{2})}\sqrt{\frac{8\kappa\alpha}{3}}$$

holds, or equivalently,

$$L\delta + L\lambda\sqrt{\alpha}\sqrt{1 - L\varrho} \le \left(\frac{\sqrt{8\kappa(1 - L\varrho)}}{\sqrt{3}(1+\sqrt{2})} - L\varrho\right)\alpha \ . \tag{126}$$

This inequation yields for α large enough as long as the coefficient of α is *positive*; it remains to show that this is always fulfilled in case (115) holds. We recall from (57) that κ was defined by

$$\kappa = 1 - \frac{3}{2}L\varrho - \gamma \ ,$$

where $\gamma < 1$ was a free parameter which has to be chosen such that κ is still positive. The coefficient of α has the structure $a - b$ with $a, b > 0$. Thus we have
$0 < a - b \Leftrightarrow 0 < (a-b)(a+b) = a^2 - b^2$ and find with (124)

$$\frac{8(1 - \frac{3}{2}L\varrho - \gamma)}{3(1+\sqrt{2})^2} - (L\varrho)^2 \ge \frac{4(1 - \frac{3}{2}L\varrho)}{3(1+\sqrt{2})^2} - (L\varrho)^2 =: p(L\varrho)$$

We observe $p(0) > 0$ and thus $p(L\varrho)$ is positive for $x_1 < L\varrho < x_2$, where $x_{1,2}$ denote the zeros of p. A simple calculation shows that for the positive zero $x_2 \approx 0.298$, and thus $p(L\varrho)$ and the coefficient of α in (126) are positive if (115) holds.

$\square$

If $x^\delta_{\alpha_{k-1}}$, $x^\delta_{\alpha_k}$ denote minimizers of the Tikhonov functional with parameters $\alpha_k = q\alpha_{k-1}$, $q < 1$, then we will show that $x^\delta_{\alpha_{k-1}} \in K_{r(\alpha_k)}(x^\delta_{\alpha_k})$ holds if only q is chosen close to 1. This means especially that $x^\delta_{\alpha_{k-1}}$ can be used as a starting value for the steepest descent method for minimizing $\Phi_{\alpha_k}(x)$.

Proposition 7.2 *Let $\alpha_k = q\alpha_{k-1}$, $\alpha_k = \delta/\eta$, $\eta \leq \varrho$, and assume that $L\varrho \leq 0.278$ holds. Then $q < 1$ can be chosen such that*

$$p = \frac{L}{q}\left[\eta + 2q\varrho + \frac{\sqrt{q}(\eta+\varrho) + \eta + q\varrho}{4\sqrt{1-\varrho L}}\right] \leq \tilde{p} < 1 \tag{127}$$

and, with $\alpha_{min} = \delta/\varrho$,

$$\frac{(\eta+2\varrho)|1/q-1|}{2(1-\tilde{p})}\max\{1, 1/\sqrt{\alpha_{min}}\} \leq \frac{1}{2L}\min\left\{\frac{1}{1+\sqrt{2}}\sqrt{\frac{8\kappa}{3}}, \frac{8\kappa}{3K(2+\sqrt{8})}\right\} \tag{128}$$

holds (for the definition of κ, K cf. (57), (42)). Moreover, $x^\delta_{\alpha_{k-1}} \in K_{r(\alpha_k)}(x^\delta_{\alpha_k})$.

<u>Proof:</u>
According to Proposition 6.3 condition (99) is fulfilled, and for $q \to 1$ converges p to the left hand side of (99), and $\tilde{p} < 1$ can be found such that (127) holds for $q < 1$ big enough. The left hand side of (128) converges to zero for $q \to 1$, and therefore q can be chosen such that both (127), (128) hold. Now let q be chosen with (127), (128) . According to (100) yields with $\alpha_k = q\alpha_{k-1}$

$$\begin{aligned}\|x^\delta_{\alpha_k} - x^\delta_{\alpha_{k-1}}\| &\leq \frac{|\alpha_{k-1}-\alpha_k|(\eta+2\varrho)}{2\sqrt{q\alpha_{k-1}}(1-p)} \\ &\leq \frac{(\eta+2\varrho)|1/q-1|}{2\sqrt{\alpha_k}(1-\tilde{p})}\alpha_k\ .\end{aligned} \tag{129}$$

Because of $\alpha_k = \delta/\eta \geq \alpha_{min}$ follows

$$\|x^\delta_{\alpha_k} - x^\delta_{\alpha_{k-1}}\| \leq \frac{(\eta+2\varrho)|1/q-1|}{2\sqrt{\alpha_{min}}(1-\tilde{p})}\alpha_k\ . \tag{130}$$

On the other hand, we conclude from (129)

$$\|x^\delta_{\alpha_k} - x^\delta_{\alpha_{k-1}}\| \leq \frac{(\eta+2\varrho)|1/q-1|}{2(1-\tilde{p})}\sqrt{\alpha_k}\ , \tag{131}$$

and combining (130), (131) yields

$$\|x^\delta_{\alpha_k} - x^\delta_{\alpha_{k-1}}\| \leq \frac{(\eta+2\varrho)|1/q-1|}{2(1-\tilde{p})}\cdot\max\left\{\frac{1}{\sqrt{\alpha_{min}}}, 1\right\}\cdot\min\{\alpha_k, \sqrt{\alpha_k}\}\ . \tag{132}$$

According to Proposition 4.5 is $x^\delta_{\alpha_{k-1}} \in K_{r(\alpha_k)}(x^\delta_{\alpha_k})$ if $\|x^\delta_{\alpha_{k-1}} - x^\delta_{\alpha_k}\| \le r(\alpha_k)$ holds. We find

$$\begin{aligned} r(\alpha_k) &= \frac{1}{L}\min\left\{\frac{1}{1+\sqrt{2}}\sqrt{\frac{8\kappa\alpha_k}{3}}, \frac{8\kappa\alpha_k}{3K(2+\sqrt{8})}\right\} \\ &\ge \frac{1}{L}\min\left\{\frac{1}{1+\sqrt{2}}\sqrt{\frac{8\kappa}{3}}, \frac{8\kappa}{3K(2+\sqrt{8})}\right\}\min\{\alpha_k, \sqrt{\alpha_k}\}\ . \qquad (133) \end{aligned}$$

Altogether we have shown

$$\begin{aligned} \|x^\delta_{\alpha_k} - x^\delta_{\alpha_{k-1}}\| &\overset{(132)}{\le} \frac{(\eta+2\varrho)|1/q-1|}{2(1-\tilde{p})}\cdot\max\left\{\frac{1}{\sqrt{\alpha_{min}}}, 1\right\}\cdot\min\{\alpha_k, \sqrt{\alpha_k}\} \\ &\overset{(128)}{\le} \frac{1}{2L}\min\left\{\frac{1}{1+\sqrt{2}}\sqrt{\frac{8\kappa}{3}}, \frac{8\kappa}{3K(2+\sqrt{8})}\right\}\min\{\alpha_k, \sqrt{\alpha_k}\} \qquad (134) \\ &\overset{(133)}{\le} \frac{1}{2}r(\alpha_k) \qquad (135) \end{aligned}$$

and thus $x^\delta_{\alpha_{k-1}} \in K_{r(\alpha_k)}(x^\delta_{\alpha_k})$.

□

For the TIGRA–algorithm we do not know the minimizing functions $x^\delta_{\alpha_{k-1}}$, but we will show that the sequence $\{x^\delta_{k-1,l}\}_{l\in\mathbb{N}}$ converges to $x^\delta_{\alpha_{k-1}}$. By choosing $TOL(k)$ in (119) small enough, we can still ensure that an iterate $x^\delta_{k-1,l}$ with $\|\Phi_{\alpha_{k-1}}(x^\delta_{k-1,l})\| \le TOL(k)$ is a suitable starting value for minimizing Φ_{α_k} with the gradient method:

Proposition 7.3 *In the TIGRA–algorithm, let q be chosen according to (127), (128), γ according to (124) and assume $\alpha_k \ge \alpha_{min}$. Moreover, assume that $x^\delta_{k-1,0} \in K_{r(\alpha_{k-1})}(x^\delta_{\alpha_{k-1}})$, and that the iterates $x^\delta_{k,l}$ are computed by the inner iteration (119)–(121), where the scaling parameter $\beta_{k,l}$ in (120) is chosen as in (94). If x^δ_{k-1,l_*} is the first iterate with*

$$\|\nabla\Phi_{\alpha_{k-1}}(x^\delta_{k-1,l_*})\| \le TOL(k-1)\ , \qquad (136)$$

$$TOL(k-1) = \frac{\gamma\alpha}{2L}\min\left\{\frac{1}{1+\sqrt{2}}\sqrt{\frac{8\kappa}{3}}, \frac{8\kappa}{3K(2+\sqrt{8}}\right\}\min\{\alpha_k, \sqrt{\alpha_k}\}\ , \qquad (137)$$

then $x^\delta_{k-1,l_*} \in K_{r(\alpha_k)}(x^\delta_{\alpha_k})$.

Proof:
It is $x^\delta_{k-1,0} \in K_{r(\alpha_{k-1})}(x^\delta_{\alpha_{k-1}})$, and thus the sequence $(x^\delta_{k-,l})$ with scaling parameter $\beta_{k,l}$ chosen as in (94) converges to $x^\delta_{\alpha_{k-1}}$ (cf. Theorem 5.6), and the inner iteration terminates after a finite number l_*. The iteration error can be estimated by (95). We conclude

$$\|x^\delta_{\alpha_k} - x^\delta_{k-1,l_*}\| \quad\le\quad \|x^\delta_{\alpha_k} - x^\delta_{\alpha_{k-1}}\| + \|x^\delta_{\alpha_{k-1}} - x^\delta_{k-1,l_*}\|$$

$$\overset{(95)}{\leq} \quad \|x^\delta_{\alpha_k} - x^\delta_{\alpha_{k-1}}\| + \frac{1}{\gamma\alpha}\|\nabla\Phi_{\alpha_{k-1}}(x^\delta_{k-1,l_*})\|$$

$$\overset{(136)}{\leq} \quad \|x^\delta_{\alpha_k} - x^\delta_{\alpha_{k-1}}\| + \frac{1}{\gamma\alpha}TOL(k-1)$$

$$\overset{(134),(137)}{\leq} \quad \frac{1}{L}\min\left\{\frac{1}{1+\sqrt{2}}\sqrt{\frac{8\kappa}{3}}, \frac{8\kappa}{3K(2+\sqrt{8})}\right\}\min\{\alpha_k, \sqrt{\alpha_k}\}$$

$$\overset{(133)}{\leq} \quad r(\alpha_k) \ ,$$

and thus $x^\delta_{k-1,l_*} \in K_{r(\alpha_k)}(x^\delta_{\alpha_k})$.

□

Next, we will show that the TIGRA–algorithm terminates after a finite number of iteration steps. Because the previous Propositions were based on the assumption that $\alpha_k \geq \alpha_{min}$ holds, we will additionally show that the terminal regularization parameter is not smaller than α_{min}.

Proposition 7.4 *Let the assumptions of Propositions 7.1–7.3 hold. For given x_0 with (123), $\bar{x}$ with (114) choose α_0 such that $x_0 \in K_{r(\alpha_0)}(x^\delta_{\alpha_0})$. Moreover, assume*

$$q \geq \frac{2}{3} \tag{138}$$

and that $TOL(k)$ admits the inequation

$$\frac{K}{\gamma\alpha}TOL(k) + \frac{L}{2(\gamma\alpha)^2}TOL(k)^2 \leq \delta \tag{139}$$

for all $k \in \mathbb{N}$. Then the TIGRA–algorithm terminates after a finite number k_* of outer iteration steps, and the last regularization parameter fulfills $\alpha_{k_*} \geq \alpha_{min}$.

Proof:
Following Proposition 7.1, we choose α_0 such that (126) holds, and find $x_0 \in K_{r(\alpha_0)}(x^\delta_{\alpha_0})$. Without loss of generality we assume $\alpha_0 > \alpha_{min}$ and because of (114) α_0 can be chosen large enough such that

$$\|y^\delta - F(x^\delta_{\alpha_0})\| \geq 5\delta$$

holds (cf. proof of Conclusion 6.2). Theorem 5.6 shows that $x_{0,l}$ converges towards $x^\delta_{\alpha_0}$, and because of $\nabla\Phi_{\alpha_0}(x^\delta_{0,l}) \to 0$ for $l \to \infty$, the inner iteration terminates after a finite number $l_*(0)$ of iteration steps. If q and $TOL(0)$ admit (127), (128), (137), then $x^\delta_{0,l_*(0)} \in K_{r(\alpha_1)}(x^\delta_{\alpha_1})$. By induction, we find for all $k \in \mathbb{N}$ with $\alpha_k = q\alpha_{k-1}$ and $\alpha_k \geq \alpha_{min}$ that the inner iteration terminates and that $x^\delta_{k-1,l_*(k-1)} \in K_{r(\alpha_k)}(x^\delta_{\alpha_k})$.
Because of $q < 1$, the sequence $\{\alpha_k\}_{k\in\mathbb{N}}$ is monotone decreasing and converges to zero. Now assume that the outer iteration does not stop or stops with an regularization parameter smaller than α_{min}. Then there exists $k \in \mathbb{N}$ with $\alpha_k = q\alpha_{k-1} < \alpha_{min}$ and $\alpha_{k-1} \geq \alpha_{min}$. The iterates $x^\delta_{k-1,l}$ converge towards $x^\delta_{\alpha_{k-1}}$. Let us denote by $l_*(k-1)$ the stopping index of the inner iteration

for the $(k-1)th$ outer iteration; for simplicity of notation, we will denote the belonging iterate by x^{δ}_{k-1,l_*}. We have by (16) and (17)

$$\begin{aligned}
\|F(x^{\delta}_{k-1,l_*}) - F(x^{\delta}_{\alpha_{k-1}})\| &\leq \|F'(x^{\delta}_{\alpha_{k-1}})\|\|x^{\delta}_{k-1,l_*} - x^{\delta}_{\alpha_{k-1}}\| + \frac{L}{2}\|x^{\delta}_{k-1,l_*} - x^{\delta}_{\alpha_{k-1}}\|^2 \\
&\overset{(40)}{\leq} K\|x^{\delta}_{k-1,l_*} - x^{\delta}_{\alpha_{k-1}}\| + \frac{L}{2}\|x^{\delta}_{k-1,l_*} - x^{\delta}_{\alpha_{k-1}}\|^2 && (140) \\
&\overset{(95),(136)}{\leq} \frac{K}{\gamma\alpha}TOL(k-1) + \frac{L}{2(\gamma\alpha)^2}TOL(k-1)^2 \\
&\overset{(139)}{\leq} \delta && (141)
\end{aligned}$$

and thus by the definition of $\alpha_{min} = \delta/\varrho$, $\|\omega\| \leq \varrho$ and $q \geq 2/3$,

$$\begin{aligned}
\|y^\delta - F(x^{\delta}_{k-1,l_*})\| &\leq \|y^\delta - F(x^{\delta}_{\alpha_{k-1}})\| + \|F(x^{\delta}_{\alpha_{k-1}}) - F(x^{\delta}_{k-1,l_*})\| \\
&\overset{(11),(141)}{\leq} \delta + 2\alpha_{k-1}\|\omega\| + \delta \\
&= \delta + 2\frac{\alpha_k}{q}\|\omega\| + \delta \\
&\leq \delta + 2\frac{\alpha_{min}}{q}\|\omega\| + \delta \\
&\leq \delta + 2\frac{\delta}{q\varrho}\|\omega\| + \delta \\
&\leq 5\delta \ ,
\end{aligned}$$

which means that the iteration would have terminated with x_{k-1,l_*} and $\alpha_{k-1} \geq \alpha_{min}$.

□

It now remains to give a convergence rate result for the TIGRA–algorithm.

Theorem 7.5 *Let F be a twice continuous (Fréchet–) differentiable operator with Lipschitz–continuous first derivative, fulfilling (7), and x_* be a solution of the equation $F(x) = y$ with*

$$x_* - \bar{x} = F'(x_*)^*\omega \qquad \|\omega\| \leq \varrho \leq 0.278 \ . \tag{142}$$

Additionally, assume that only noisy data y^δ with $\|y^\delta - y\| \leq \delta$ are given, that $\bar{x}$ admits (114) and let the parameters x_0, α_0, q, $TOL(k)$, γ and $\beta_{k,l}$ for the TIGRA–algorithm are chosen as follows:

1. *x_0 with (123),*
2. *α_0 with (125),*
3. *q with (127), (128), (138) and*

$$\frac{K(\eta+2\varrho)|1/q-1|}{2(1-\tilde{p})}\sqrt{\alpha_0} + \frac{L}{2}\left(\frac{(\eta+2\varrho)|1/q-1|}{2(1-\tilde{p})}\right)^2\alpha_0 \leq 3\delta \ , \tag{143}$$

4. TOL(k) with (136), (137) ,

5. γ with (124) , and

6. $\beta_{k,l}$ with (94) .

Then, if $x^{\delta}_{k_*,l_*(k_*)}$ denotes the last iterate of the TIGRA–algorithm, we get the error estimate

$$\|x^{\delta}_{k_*,l_*(k_*)} - x_*\| \leq \left(\frac{14\|\omega\|}{1-L\|\omega\|}\right)^{1/2} \delta^{1/2} + \frac{1}{K}\delta \tag{144}$$

(for the definition of K, cf. (42)). For $\delta \to 0$, we obtain esp.

$$\|x^{\delta}_{k_*,l_*(k_*)} - x_*\| = O(\delta^{1/2}) \; . \tag{145}$$

Proof:
According to Proposition 7.4 does the iteration terminate. In the following, we might denote by $l_*(k)$ iteration index where the inner iteration (119)–(121) for a fixed k terminates. Then the last two outer iterates $x^{\delta}_{k_*-1,l_*(k_*-1)}$, $x^{\delta}_{k_*,l_*(k_*)}$ fulfill

$$\begin{aligned} \|y^{\delta} - F(x^{\delta}_{k_*,l_*(k_*)})\| &\leq 5\delta \\ \|y^{\delta} - F(x^{\delta}_{k_*-1,l_*(k_*-1)})\| &> 5\delta \; , \end{aligned} \tag{146}$$

and therefore

$$\begin{aligned} 5\delta &< \|y^{\delta} - F(x^{\delta}_{k_*-1,l_*(k_*-1)})\| \leq \|y^{\delta} - F(x^{\delta}_{\alpha_{k_*-1}})\| + \|F(x^{\delta}_{\alpha_{k_*-1}}) - F(x^{\delta}_{k_*-1,l_*(k_*-1)})\| \\ &\overset{(141)}{\leq} \|y^{\delta} - F(x^{\delta}_{\alpha_{k_*-1}})\| + \delta \; , \end{aligned}$$

i.e.

$$4\delta \leq \|y^{\delta} - F(x^{\delta}_{\alpha_{k_*-1}})\| \; . \tag{147}$$

Now let us assume that $\|y^{\delta} - F(x^{\delta}_{\alpha_{k_*}})\| < \delta$ holds. Then

$$\begin{aligned} 3\delta &< \|y^{\delta} - F(x^{\delta}_{\alpha_{k_*-1}})\| - \|y^{\delta} - F(x^{\delta}_{\alpha_{k_*}})\| \\ &\leq \|F(x^{\delta}_{\alpha_{k_*-1}}) - F(x^{\delta}_{\alpha_{k_*}})\| \\ &\leq K\|x^{\delta}_{\alpha_{k_*-1}} - x^{\delta}_{\alpha_{k_*}}\| + \frac{L}{2}\|x^{\delta}_{\alpha_{k_*-1}} - x^{\delta}_{\alpha_{k_*}}\|^2 \; . \end{aligned}$$

Here, the last inequation was obtained similarly to (140). By (129) yields

$$\begin{aligned} \|x^{\delta}_{\alpha_{k_*}} - x^{\delta}_{\alpha_{k_*-1}}\| &\leq \frac{(\eta+2\varrho)|1/q-1|}{2(1-\tilde{p})}\sqrt{\alpha_{k_*}} \\ &\leq \frac{(\eta+2\varrho)|1/q-1|}{2(1-\tilde{p})}\sqrt{\alpha_0} \end{aligned} \tag{148}$$

and thus

$$3\delta \;<\; \frac{K(\eta+2\varrho)|1/q-1|}{2(1-\tilde{p})}\sqrt{\alpha_0}+\frac{L}{2}\left(\frac{(\eta+2\varrho)|1/q-1|}{2(1-\tilde{p})}\right)^2\alpha_0$$

This is obviously a contradiction if q is chosen with (143) and we have shown

$$\begin{aligned}\delta \;&\leq\; \|y^\delta - F(x^\delta_{\alpha_{k_*}})\| \leq \|y^\delta - F(x^\delta_{k_*,l_*(k_*)})\| + \|F(x^\delta_{k_*,l_*(k_*)}) - F(x^\delta_{\alpha_{k_*}})\| \\ &\overset{(146),(141)}{\leq}\; 6\delta \end{aligned} \tag{149}$$

Thus, $x^\delta_{\alpha_{k_*}}$ fulfills the modified version (14) of Morozov's discrepancy principle, and , by using (149), (15) we have therefore shown

$$\|x^\delta_{\alpha_{k_*}} - x_*\| \leq \left(\frac{14\|\omega\|}{1-L\|\omega\|}\right)^{1/2}\delta^{1/2}\ ,$$

and conclude finally

$$\begin{aligned}\|x^\delta_{k_*,l_*(k_*)} - x_*\| \;&\leq\; \|x^\delta_{k_*,l_*(k_*)} - x^\delta_{\alpha_{k_*}}\| + \|x^\delta_{\alpha_{k_*}} - x_*\| \\ &\overset{(95),(136)}{\leq}\; \frac{1}{\gamma}TOL(k_*) + \left(\frac{14\|\omega\|}{1-L\|\omega\|}\right)^{1/2}\delta^{1/2} \\ &\overset{(139)}{\leq}\; \frac{1}{K}\delta + \left(\frac{14\|\omega\|}{1-L\|\omega\|}\right)^{1/2}\delta^{1/2}\end{aligned}$$

□

The classical convergence rate result for Tikhonov–regularization was given by Engl, Kunisch and Neubauer [10] for Fréchet differentiable operators with Lipschitz continuous derivative and a smoothness condition as above with $L\varrho < 1$. But it was then an open question how the minimizers of the Tikhonov–functional could be computed. Here we have presented an algorithm which actually computes an approximation to a minimizer of the functional that fulfills Morozov's discrepancy principle. The used conditions– twice Fréchet differentiable operator with Lipschitz–continuous derivative and $L\varrho \leq 0.278$ – are only slightly stronger than in the above mentioned paper.

8 Applications

Within this section we might give some practical relevant examples that meet the requirements of the TIGRA–algorithm. The main focus will be on SPECT (Single Photon Emission Computed Tomography), a medical imaging technology.

Bilinear Operator Equations

Many operators $F : X_1 \times X_2 \to Y$ can be decomposed into (or approximated by) a sum of a continuous linear operator A and a bilinear operator B ,

$$F(x) = Af + B(f, \mu) \ , \tag{150}$$

$x = (f, \mu)$. These operators have been extensively treated in [25]. If an estimate

$$\|B(f, \mu)\| \le \|B\| \|f\| \|\mu\|$$

holds,

$$\|B\| = \inf_{c \in \mathbb{R}^+} \{ \|B(f, \mu)\| \le c \|f\| \|\mu\| \mid (f, \mu) \in X_1 \times X_2 \},$$

then it is easy to see that F is twice continuous Fréchet–differentiable with Lipschitz–continuous first derivative, and the TIGRA–algorithm can therefore be used. A classical example for a bilinear operator is the *autoconvolution* operator,

$$F(x)(s) = \int_a^b x(s - t) x(t) \, dt \ , \qquad -\infty < a < b < \infty \ ,$$

cf. [14, 24]. Other examples appear in parameter estimation problems for partial differential equations [3, 8, 16, 22], and in [26], a bilinear approximation to the attenuated Radon transform was used.

Single Photon Emission Computed Tomography

In SECT, the aim is to reconstruct the distribution of a radiopharmaceutical inside a (human) body from measurements of the radiation outside the body. The connection between the measurements y and the activity distribution f is given by the attenuated Radon transform:

$$y = R(f, \mu)(s, \omega) = \int_{\mathbb{R}} f(s\omega^\perp + t\omega) e^{-\int_t^\infty \mu(s\omega^\perp + \tau\omega) \, d\tau} \, dt \ , \tag{151}$$

$s \in \mathbb{R}, \ \omega \in S^1$. As for the Radon transform, the data are represented as line integrals over all possible unit vectors ω. The (usually also unknown) function μ is called the attenuation map, it is related to the density of the tissue and reflects the fact that the intensity of the emitted γ–rays is damped when traveling through the body. In general the attenuation map will be known only if an additional CT (Computerized Tomography)–scan is performed, which might cause a significant rise of costs for the medical examination. It is therefore of interest to reconstruct f without knowledge of μ, i.e. to solve the nonlinear problem (151) [19, 20, 31, 21, 26, 25]. Dicken [7] studied the mapping properties of the attenuated Radon transform. He also introduced a redefined attenuated Radon transform R_ϱ by

$$R_\varrho(f, \mu)(s, \omega) = \int_{\mathbb{R}} f(s\omega^\perp + t\omega) E \left[\int_{-\rho}^{\rho} \mu(s\omega^\perp + \tau\omega) \, d\tau \right] dt \ . \tag{152}$$

The function $E \in C^2(\mathbb{R})$ is chosen such that

$$E(x) = exp(-x) \text{ for } x \in \mathbb{R}^+$$

and $|E|$, $|E'|$ and $|E''|$ are bounded. For SPECT, the functions f and μ will be nonnegative with compact support. If we assume that f has its support in a disc with radius ϱ, then the operator R_ϱ coincides with R for admissible sets (f, μ). Fortunately, R_ϱ has much better mapping properties than R. If R_ϱ is considered as operator

$$R_\varrho : H_0^{s_1}(\Omega) \times H_0^{s_2}(\Omega) \to L_2(S^1 \times [-\varrho, \varrho]) \qquad s_1, s_2 > 0 \ ,$$

where $H_0^s(\Omega)$ denotes a Sobolev space with zero boundary conditions, then it can be shown that R_ϱ is twice continuous Fréchet–differentiable with Lipschitz–continuous first derivative whenever (s_1, s_2) are chosen with

$$2 < \frac{3s_1}{2(1-s_2)} \ , \tag{153}$$

cf. [7] . The function μ is related to the density of the tissue and has therefore discontinuities, and thus $s_2 < 1/2$ is a realistic assumption. If we set $s_2 = 1/2 - \varepsilon$, $\varepsilon > 0$, then a choice

$$s_1 > \frac{2(1+2\varepsilon)}{3} > \frac{2}{3}$$

ensures the above given properties of R_ϱ. However, for practical applicability we would like to have (s_1, s_2) as small as possible. If we additionally assume that f is bounded, then we can define

$$D(R_\varrho) := D_{s_1,s_2,C} = \{(f, \mu) \in H_0^{s_1}(\Omega) \times H_0^{s_2}(\Omega) \mid \|f\|_\infty \le C\} \ , \tag{154}$$

and consider

$$R_\varrho : D_{s_1,s_2,C} \to L_2(S^1 \times [-\varrho, \varrho]) \qquad s_1, s_2 > 0 \ .$$

With this definition area, R_ϱ is twice continuous Fréchet–differentiable with Lipschitz–continuous first derivative for

$$s_2 > \frac{1}{4} \ , \qquad s_1 > \frac{2}{3}(1 - s_2) \ ,$$

cf. [6]. Thus both s_1, s_2 can be chosen smaller than 1/2, e.g. $s_1 > 4/9$, $s_2 = 1/3$. Altogether we have seen that the redefined attenuated Radon transform fulfills the requirements of our TIGRA–algorithm.

For our numerical test computations we used the so called MCAT–phantom [29]. The activity f_* is concentrated in the heart, and the attenuation function μ_* models a cut through the thorax.

The generated data y were blurred with random noise (relative error $\delta_{rel} \in \{1\%, 2.5\%, 5\%, 10\%\}$), and the TIGRA–algorithm was used to reconstruct the activity and attenuation function. For all four reconstructions $\alpha_0 = 10000$, $q = 0.7$ and $\bar{x} = (\bar{f}, \bar{\mu}) = (0, 0)$ was used, and the iteration stopped always with a regularization parameter $\alpha_{k_*(\delta)}$ with $\delta \le \|y^\delta - R_\varrho(f^\delta_{k_*(\delta),l_*(\delta)}, \mu^\delta_{k_*(\delta),l_*(\delta)})\| \le 5\delta$ (cf. Table 3).

Comparing the reconstructions for the activity and attenuation functions with the original functions (f, μ) in Figure 1, then we find that the reconstructions for the activity function look quite good, but the reconstruction for the attenuation map is far off. This is a well known phenomena for SPECT with both (f, μ) unknown. Because of the nonlinearity of R_ϱ, the solution of (151) will not be unique. Tikhonov–regularization will then give an approximation

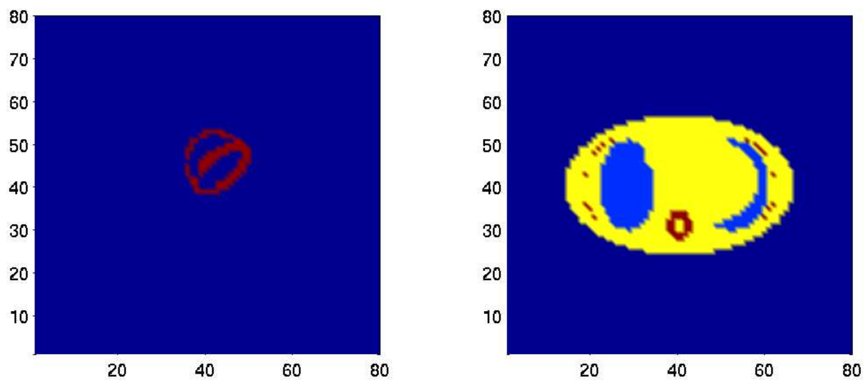

Figure 1: Activity function f_* (left) and attenuation function μ_* (right)

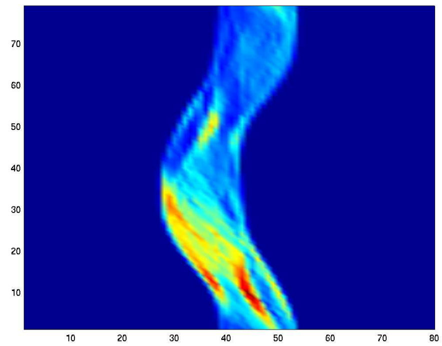

Figure 2: Generated data y.

Figure 3: Residuum and regularization parameter $\alpha_{k_*(\delta)}$ for the terminal iterate of the TIGRA–algorithm

δ_{rel}	δ	$\|y^\delta - R_\varrho(f^\delta_{k_*(\delta),l_*(\delta)}, \mu^\delta_{k_*(\delta),l_*(\delta)})\|$	5δ	$\alpha_{k_*(\delta)}$
10%	0.6601	3.256	3.3	403.5
5%	0.33	1.582	1.65	23.26
2.5%	0.2063	1.026	1.032	7.98
1%	0.0826	0.402	0.413	0.322

to a solution of (151) with minimal distance to the a priori guess $\bar{x} = (\bar{f}, \bar{\mu}) = (0, 0)$. This solution does not necessarily coincide with the solution given in Figure 1. If one would like to have a better reconstruction for the attenuation function, then one has to come up with a better a priori guess $\bar{\mu}$. However, in SPECT we are only interested in reconstructing the activity function by using an attenuation map that is consistent to the data. Because our reconstruction for the activity function differs from f_* by a factor, we measured the reconstruction quality by

$$reconstruction\ accuracy = \left\| \frac{f^\delta_{k_*(\delta),l_*(\delta)}}{\|f^\delta_{k_*(\delta),l_*(\delta)}\|} - \frac{f_*}{\|f_*\|} \right\| \cdot 100\% \ .$$

The results are given in Table 4. Visual examination (Figure 5) shows that the image gets sharper for smaller noise and that especially the south–western part of the activity function is better reconstructed. Our numerical test computations show that TIGRA can be used in order to reconstruct at least a good approximation to the activity function f_*. We might especially remark that the algorithm seems to pick a good final regularization parameter $\alpha_{k_*(\delta)}$.

Figure 4: Reconstruction accuracy of the TIGRA–algorithm with blurred data

relative data noise	reconstruction accuracy
10%	54%
5%	25%
2.5%	20%
1%	18%

References

[1] A. W. Bakushinskii. The problem of the convergence of the iteratively regularized Gauss–Newton method. *Comput. Maths. Math. Phys.*, (32):1353–1359, 1992.

[2] B. Blaschke, A. Neubauer, and O. Scherzer. On convergence rates for the iteratively regularized Gauss–Newton method. *IMA Journal of Numerical Analysis*, (17):421–436, 1997.

[3] D. Colton and M. Piana. The simple method for solving the electromagnetic inverse scattering problem: the case of TE polarized waves. *Inverse Problems*, 14(3):597–614, 1998.

[4] C. Cravaris and J. H. Seinfeld. Identification of parameters in distributed parameter systems by regularization. *SIAM J. Contr. Opt.*, (23):217–241, 1985.

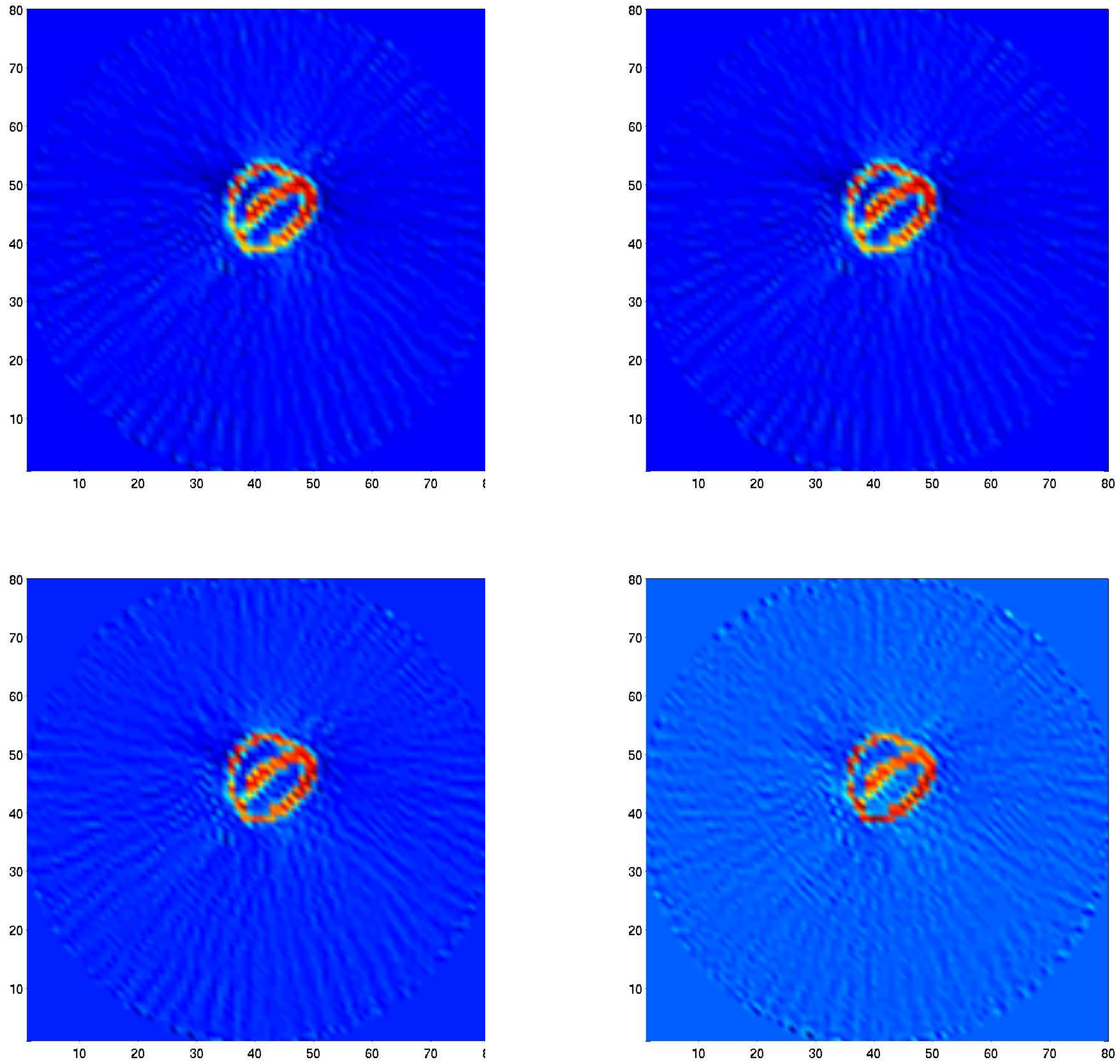

Figure 5: Reconstruction of the activity function from noisy data with 10% (top left), 5% (top right), 2.5% (bottom left) and 1% (bottom right) relative noise.

[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] V. Dicken. *Simultaneous Activity and Attenuation Reconstruction in Single Photon Emission Computed Tomography, a Nonlinear Ill-Posed Problem.* Ph.d. thesis, Universität Potsdam, 5/ 1998.

[7] V. Dicken. A new approach towards simultaneous activity and attenuation reconstruction in emission tomography. *Inverse Problems*, 15(4):931–960, 1999.

[8] O. Dorn. A transport-backtransport method for optical tomography. *Inverse Problems*, 14(5):1107–1130, 1998.

[9] H. W. Engl, M. Hanke, and A. Neubauer. *Regularization of Inverse Problems.* Kluwer, Dordrecht, 1996.

[10] H.W. Engl, K. Kunisch, and A. Neubauer. Convergence rates for Tikhonov regularization of nonlinear ill-posed problems. *Inverse Problems*, (5):523–540, 1989.

[11] M. Hanke. A regularizing Levenberg–Marquardt scheme, with applications to inverse groundwater filtration problems. *Inverse Problems*, (13):79–95, 1997.

[12] M. Hanke. Regularizing properties of a truncated Newton–cg algorithm for nonlinear ill-posed problems. *Num. Funct. Anal. Optim.*, (18):971–993, 1997.

[13] M. Hanke, A. Neubauer, and O. Scherzer. A convergence analysis of the Landweber iteration for nonlinear ill-posed problems. *Numerische Mathematik*, (72):21–37, 1995.

[14] B. Hofmann and R. Gorenflo. On autoconvolution and regularization. *Inverse Problems*, (10):353–373, 1994.

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

[16] T. Klibanov, T.R. Lucas, and R.M. Frank. A fast and accurate imaging algorithm in optical/diffusion tomography. *Inverse Problems*, 13(5):1341–1363, 1997.

[17] D. Luenberger. *Optimization by Vector Space Methods.* Wiley, New York, 1969.

[18] V. A. Morozov. *Methods for Solving Incorrectly Posed Problems.* Springer, New York, 1984.

[19] F. Natterer. Computerized tomography with unknown sources. *SIAM J. Appl. Math.*, (43):1201–12, 1983.

[20] F. Natterer. Determination of tissue attenuation in emission tomography of optically dense media. *Inverse Problems*, (9):731–736, 1993.

[21] F. Natterer. Inversion of the attenuated radon transform. *Inverse Problems*, 17(1):113–119, 2001.

[22] F. Natterer and F. Wubbeling. A propagation-backpropagation method for ultrasound tomography. *Inverse Problems*, 11(6):1225–1232, 1998.

[23] R. Ramlau. A modified Landweber–method for inverse problems. *Numerical Functional Analysis and Optimization*, 20(1& 2), 1999.

[24] R. Ramlau. Morozov's discrepancy principle for Tikhonov regularization of nonlinear operators. *Numer. Funct. Anal. and Optimiz*, 23(1&2):147–172, 2002.

[25] R. Ramlau. A steepest descent algorithm for the global minimization of the Tikhonov–functional. *Inverse Problems*, 18(2):381–405, 2002.

[26] R. Ramlau, R. Clackdoyle, F. Noo, and G. Bal. Accurate attenuation correction in SPECT imaging using optimization of bilinear functions and assuming an unknown spatially–varying attenuation distribution. *Z. angew. Math. Mech.*, 80(9):613–621, 2000.

[27] O. Scherzer. The use of Morozov's discrepancy principle for Tikhonov regularization for solving nonlinear ill–posed problems. *Computing*, (51):45–60, 1993.

[28] O. Scherzer. A modified landweber iteration for solving parameter estimation problems. *Appl. Math. Optim.*, 38:45–68, 1998.

[29] J. A. Terry, B. M. W. Tsui, J. R. Perry, J. L. Hendricks, and G. T. Gullberg. The design of a mathematical phantom of the upper human torso for use in 3-d spect imaging research. In *Proc. 1990 Fall Meeting Biomed. Eng. Soc. (Blacksburg, VA)*, pages 1467–74. New York University Press, 1990.

[30] A.N. Tikhonov and V.Y. Arsenin. *Solutions of ill posed problems.* Winston Wiley, New York, 1977.

[31] A Welch, R Clack, F Natterer, and G T Gullberg. Toward accurate attenuation correction in SPECT without transmission measurements. *IEEE Trans. Med. Imaging*, (16):532–40, 1997.

[32] E. Zeidler. *Nonlinear Functional Analysis and its Applications.* Springer, New York, 1985.

Berichte aus der Technomathematik

ISSN 1435-7968

http://www.math.uni-bremen.de/zetem/berichte.html

— Vertrieb durch den Autor —

Reports **Stand: 3. September 2002**

98–01. Peter Benner, Heike Faßbender:
An Implicitly Restarted Symplectic Lanczos Method for the Symplectic Eigenvalue Problem, Juli 1998.

98–02. Heike Faßbender:
Sliding Window Schemes for Discrete Least-Squares Approximation by Trigonometric Polynomials, Juli 1998.

98–03. Peter Benner, Maribel Castillo, Enrique S. Quintana-Ortí:
Parallel Partial Stabilizing Algorithms for Large Linear Control Systems, Juli 1998.

98–04. Peter Benner:
Computational Methods for Linear–Quadratic Optimization, August 1998.

98–05. Peter Benner, Ralph Byers, Enrique S. Quintana-Ortí, Gregorio Quintana-Ortí:
Solving Algebraic Riccati Equations on Parallel Computers Using Newton's Method with Exact Line Search, August 1998.

98–06. Lars Grüne, Fabian Wirth:
On the rate of convergence of infinite horizon discounted optimal value functions, November 1998.

98–07. Peter Benner, Volker Mehrmann, Hongguo Xu:
A Note on the Numerical Solution of Complex Hamiltonian and Skew-Hamiltonian Eigenvalue Problems, November 1998.

98–08. Eberhard Bänsch, Burkhard Höhn:
Numerical simulation of a silicon floating zone with a free capillary surface, Dezember 1998.

99–01. Heike Faßbender:
The Parameterized SR Algorithm for Symplectic (Butterfly) Matrices, Februar 1999.

99–02. Heike Faßbender:
Error Analysis of the symplectic Lanczos Method for the symplectic Eigenvalue Problem, März 1999.

99–03. Eberhard Bänsch, Alfred Schmidt:
Simulation of dendritic crystal growth with thermal convection, März 1999.

99–04. Eberhard Bänsch:
Finite element discretization of the Navier-Stokes equations with a free capillary surface, März 1999.

99–05. Peter Benner:
Mathematik in der Berufspraxis, Juli 1999.

99–06. Andrew D.B. Paice, Fabian R. Wirth:
Robustness of nonlinear systems and their domains of attraction, August 1999.

99–07. Peter Benner, Enrique S. Quintana-Ortí, Gregorio Quintana-Ortí:
Balanced Truncation Model Reduction of Large-Scale Dense Systems on Parallel Computers, September 1999.

99–08. Ronald Stöver:
Collocation methods for solving linear differential-algebraic boundary value problems, September 1999.

99–09. Huseyin Akcay:
Modelling with Orthonormal Basis Functions, September 1999.

99–10. Heike Faßbender, D. Steven Mackey, Niloufer Mackey:
Hamilton and Jacobi come full circle: Jacobi algorithms for structured Hamiltonian eigenproblems, Oktober 1999.

99–11. Peter Benner, Vincente Hernández, Antonio Pastor:
On the Kleinman Iteration for Nonstabilizable System, Oktober 1999.

99–12. Peter Benner, Heike Faßbender:
A Hybrid Method for the Numerical Solution of Discrete-Time Algebraic Riccati Equations, November 1999.

99–13. Peter Benner, Enrique S. Quintana-Ortí, Gregorio Quintana-Ortí:
Numerical Solution of Schur Stable Linear Matrix Equations on Multicomputers, November 1999.

99–14. Eberhard Bänsch, Karol Mikula:
Adaptivity in 3D Image Processing, Dezember 1999.

00–01. Peter Benner, Volker Mehrmann, Hongguo Xu:
Perturbation Analysis for the Eigenvalue Problem of a Formal Product of Matrices, Januar 2000.

00–02. Ziping Huang:
Finite Element Method for Mixed Problems with Penalty, Januar 2000.

00–03. Gianfrancesco Martinico:
Recursive mesh refinement in 3D, Februar 2000.

00–04. Eberhard Bänsch, Christoph Egbers, Oliver Meincke, Nicoleta Scurtu:
Taylor-Couette System with Asymmetric Boundary Conditions, Februar 2000.

00–05. Peter Benner:
Symplectic Balancing of Hamiltonian Matrices, Februar 2000.

00–06. Fabio Camilli, Lars Grüne, Fabian Wirth:
A regularization of Zubov's equation for robust domains of attraction, März 2000.

00–07. Michael Wolff, Eberhard Bänsch, Michael Böhm, Dominic Davis:
Modellierung der Abkühlung von Stahlbrammen, März 2000.

00–08. Stephan Dahlke, Peter Maaß, Gerd Teschke:
Interpolating Scaling Functions with Duals, April 2000.

00–09. Jochen Behrens, Fabian Wirth:
A globalization procedure for locally stabilizing controllers, Mai 2000.

00–10. Peter Maaß, Gerd Teschke, Werner Willmann, Günter Wollmann:
Detection and Classification of Material Attributes – A Practical Application of Wavelet Analysis, Mai 2000.

00–11. Stefan Boschert, Alfred Schmidt, Kunibert G. Siebert, Eberhard Bänsch, Klaus-Werner Benz, Gerhard Dziuk, Thomas Kaiser:
Simulation of Industrial Crystal Growth by the Vertical Bridgman Method, Mai 2000.

00–12. Volker Lehmann, Gerd Teschke:
Wavelet Based Methods for Improved Wind Profiler Signal Processing, Mai 2000.

00–13. Stephan Dahlke, Peter Maass:
A Note on Interpolating Scaling Functions, August 2000.

00–14. Ronny Ramlau, Rolf Clackdoyle, Frédéric Noo, Girish Bal:
Accurate Attenuation Correction in SPECT Imaging using Optimization of Bilinear Functions and Assuming an Unknown Spatially-Varying Attenuation Distribution, September 2000.

00–15. Peter Kunkel, Ronald Stöver:
Symmetric collocation methods for linear differential-algebraic boundary value problems, September 2000.

00–16. Fabian Wirth:
The generalized spectral radius and extremal norms, Oktober 2000.

00–17. Frank Stenger, Ahmad Reza Naghsh-Nilchi, Jenny Niebsch, Ronny Ramlau:
A unified approach to the approximate solution of PDE, November 2000.

00–18. Peter Benner, Enrique S. Quintana-Ortí, Gregorio Quintana-Ortí:
Parallel algorithms for model reduction of discrete–time systems, Dezember 2000.

00–19. Ronny Ramlau:
A steepest descent algorithm for the global minimization of Tikhonov–Phillips functional, Dezember 2000.

01–01. Efficient methods in hyperthermia treatment planning:
Torsten Köhler, Peter Maass, Peter Wust, Martin Seebass, Januar 2001.

01–02. Parallel Algorithms for LQ Optimal Control of Discrete-Time Periodic Linear Systems:
Peter Benner, Ralph Byers, Rafael Mayo, Enrique S. Quintana-Ortí, Vicente Hernández, Februar 2001.

01–03. Peter Benner, Enrique S. Quintana-Ortí, Gregorio Quintana-Ortí:
Efficient Numerical Algorithms for Balanced Stochastic Truncation, März 2001.

01–04. Peter Benner, Maribel Castillo, Enrique S. Quintana-Ortí:
Partial Stabilization of Large-Scale Discrete-Time Linear Control Systems, März 2001.

01–05. Stephan Dahlke:
Besov Regularity for Edge Singularities in Polyhedral Domains, Mai 2001.

01–06. Fabian Wirth:
A linearization principle for robustness with respect to time-varying perturbations, Mai 2001.

01–07. Stephan Dahlke, Wolfgang Dahmen, Karsten Urban:
Adaptive Wavelet Methods for Saddle Point Problems - Optimal Convergence Rates, Juli 2001.

01–08. Ronny Ramlau:
Morozov's Discrepancy Principle for Tikhonov regularization of nonlinear operators, Juli 2001.

01–09. Michael Wolff:
Einführung des Drucks für die instationären Stokes–Gleichungen mittels der Methode von Kaplan, Juli 2001.

01–10. Stephan Dahlke, Peter Maaß, Gerd Teschke:
Reconstruction of Reflectivity Desities by Wavelet Transforms, August 2001.

01–11. Stephan Dahlke:
Besov Regularity for the Neumann Problem, August 2001.

01–12. Bernard Haasdonk, Mario Ohlberger, Martin Rumpf, Alfred Schmidt, Kunibert G. Siebert:
h–p–Multiresolution Visualization of Adaptive Finite Element Simulations, Oktober 2001.

01–13. Stephan Dahlke, Gabriele Steidl, Gerd Teschke:
Coorbit Spaces and Banach Frames on Homogeneous Spaces with Applications to Analyzing Functions on Spheres, August 2001.

02–01. Michael Wolff, Michael Böhm:
Zur Modellierung der Thermoelasto-Plastizität mit Phasenumwandlungen bei Stählen sowie der Umwandlungsplastizität, Februar 2002.

02–02. Stephan Dahlke, Peter Maaß:
An Outline of Adaptive Wavelet Galerkin Methods for Tikhonov Regularization of Inverse Parabolic Problems, April 2002.

02–03. Alfred Schmidt:
A Multi-Mesh Finite Element Method for Phase Field Simulations, April 2002.

02–04. Sergey N. Dachkovski, Michael Böhm:
A Note on Finite Thermoplasticity with Phase Changes, Juli 2002.

02–05. Michael Wolff, Michael Böhm:
Phasenumwandlungen und Umwandlungsplastizität bei Stählen im Konzept der Thermoelasto-Plastizität, Juli 2002.

02–06. Gerd Teschke:
Construction of Generalized Uncertainty Principles and Wavelets in Anisotropic Sobolev Spaces, August 2002.

02–07. Ronny Ramlau:
TIGRA – an iterative algorithm for regularizing nonlinear ill–posed problems, August 2002.