Zentrum für Technomathematik

Fachbereich 3 – Mathematik und Informatik

A steepest descent algorithm for the global minimization of Tikhonov–Phillips functional

Ronny Ramlau

Report 00–19

Berichte aus der Technomathematik

Report 00–19 **Dezember 2000**

A steepest descent algorithm for the global minimization of Tikhonov–Phillips functional

Ronny Ramlau
University of Bremen, Germany

January 4, 2001

Abstract

We report on a new iterative approach for finding a global minimizer of the Tikhonov–Phillips functional with a special class of nonlinear operators F. Assuming that the operator itself can be decomposed into (or approximated by) a sum of a linear and a bilinear operator, we introduce a two–step iteration scheme based on an outer iteration over the regularization parameter α and an inner iteration with a steepest descent method. Finally we present numerical results for the reconstruction of the emission function in single photon emission computed tomography (SPECT).

1 Introduction

Many problems in natural science require to solve an operator equation

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

with $F: X \to Y$ being a continuous operator between Hilbert spaces X, Y. The problem of solving (1) is called *ill posed* if its solution does not depend continuously on the data y. If we have to deal with real data, e.g. data from medical imaging, we can not hope to get exact data but data y^δ with

$$\|y - y^\delta\| \leq \delta \ . \tag{2}$$

In order to control the influence of the data error we have to employ *regularization methods*. For linear operators F, the theory of regularization is well developed. For a good overview we refer to [15, 7] . Driven by the needs of applications, the computation of solutions of *nonlinear* operator equations is getting more and more important. To this end, several of the known regularization methods for linear equations were generalized. Tikhonov regularization might be the best known method. As a solution of (1) the minimizing element of the functional

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

is taken. With appropriate parameter choice α and under slight restrictions to the operator F, it was shown in [8, 10, 9] that even for nonlinear operators the minimizing functions x_α^δ

of (3) converge to a solution x_* of (1) for $\delta \to 0$. From the numerical point of view, the difficulty in this approach is to find the global minimizer of (3). An alternative are iterative methods, which are easiest to implement. The known iteration schemes for nonlinear problems include Landweber methods [13, 20], Levenberg–Marquardt methods [11], Gauss–Newton [1, 3], conjugate gradient [12] and Newton–like methods [2]. However, the problem with all of these methods is that they impose some strong restrictions on the operator F: Firstly, it is required to have a Fréchet derivative; and secondly, certain estimates involving the operator and its Fréchet derivative have to hold. One of the widely used assumptions is

$$\|F(x) - F(\tilde{x}) - F'(x)(x - \tilde{x})\| = O(\|x - \tilde{x}\|\|F(x) - F(\tilde{x})\|) \ , \tag{4}$$

some slightly weaker assumptions are used in a recent article [5]. For practical applications, it often seems impossible to prove (4). Considering *bilinear* operator equations (to be defined later on), one will hardly be able to prove such estimates.

In this paper we will focus on Tikhonov–Phillips regularization. In this context, one usually encounters two problems: first, to use a proper regularization parameter α, and second, to *compute* the minimizing element. For linear operator equations $Ax = y$ and a given parameter α, the Tikhonov–Phillips functional is convex and the minimizer x_α^δ of the Tikhonov–Phillips functional is simply computed by

$$x_\alpha^\delta = (A^*A + \alpha I)^{-1} A^* y^\delta \ . \tag{5}$$

If α is known a priori, (5) has to be solved only once; however, by using an a posteriori parameter strategy, (5) has to be computed quite often: Starting with $\alpha_0, q > 0$, one computes $\alpha_j = \alpha_0 q^j$ and $x_{\alpha_j}^\delta$ until $\|y^\delta - A x_{\alpha_j}^\delta\| \le c\delta$ for the first time (Morozov's discrepancy principle). The focus here is to compute the minimizing functions $x_{\alpha_j}^\delta$ *fast* [18, 17]. In case of a nonlinear operator F, (3) is no longer convex and, even for simple operators, it might have several local minima. By using an iteration scheme for minimizing (3), the iterates will usually converge to a local minimum closest to the starting value of the iteration. This is illustrated with a simple example in the last section. To ensure convergence to a global minimum, we will introduce a combination of Tikhonov–Phillips regularization and a steepest descent method. The main idea of our approach comes from the observation that, for a certain class of operators, the Tikhonov–Phillips functional is still locally convex in a neighborhood of a global minimum, and that the size of this neighborhood grows with the regularization parameter α. If we want to minimize (3) for a certain parameter α the proposed algorithm works as follows: If we know an approximation $\tilde{x}$ to the solution of (1) which lies in the convex neighborhood of a minimizing element x_α^δ, the steepest descent method can be used immediately. If not, we choose a sufficiently big parameter $\alpha_0 > \alpha$, s.t. $\tilde{x}$ is in the convexity area of Φ_{α_0}, and compute $x_{\alpha_0}^\delta$ via steepest descent and starting value $\tilde{x}$. For $0 < q < 1$, we set $\alpha_1 = q^1 \alpha_0$. Under certain conditions, $x_{\alpha_0}^\delta$ is in the convexity area of Φ_{α_1} and $x_{\alpha_1}^\delta$ can be computed again with steepest descent and starting value $x_{\alpha_0}^\delta$. This process is repeated until we finally have a starting value where the steepest descent iterates converge to the minimizing element x_α^δ of Φ_α. By this way, we can ensure that the iteration will always converge to a *global* minimum of the Tikhonov–Phillips functional and not only to a local one.

For what follows we will restrict our attention to operator equations, which can be decomposed as

$$F(x) = Af + B(f,\mu) \ , \tag{6}$$

with $x = (f,\mu) \in X_1 \times X_2$, X_1, X_2 Hilbert spaces, A a continuous linear operator in f and B a bilinear operator in (f,μ):

$$\begin{aligned} A &: X_1 \to Y &(7)\\ B &: X_1 \times X_2 \to Y &(8) \end{aligned}$$

$$\begin{aligned} B(\lambda(f_1+f_2),\mu) &= \lambda(B(f_1,\mu)+B(f_2,\mu)) &(9)\\ B(f,\lambda(\mu_1+\mu_2)) &= \lambda(B(f,\mu_1)+B(f,\mu_2)) &(10)\\ \|B(f,\mu)\| &\le \|B\|\|f\|\|\mu\| \ . &(11) \end{aligned}$$

To shorten the notation, equation (6) will be scaled such that

$$\|A\| + 2\|B\| \le 1. \tag{12}$$

The inner product on $X_1 \times X_2$ is defined by

$$\langle (f_1,\mu_1),(f_2,\mu_2)\rangle := \langle f_1, f_2\rangle + \langle \mu_1,\mu_2\rangle \ .$$

For example, the convolution operator

$$T(f,g)(s) = \int f(t)g(t-s)\,dt \tag{13}$$

is bilinear; twice Fréchet differentiable operators can be well approximated by

$$F(x) \approx F(x_0) + F'(x_0)(x) + F''(x_0)(x,x) \ . \tag{14}$$

With measured data $y = F(x)$ and a known approximation x_0 to the solution, (14) can be transformed into

$$\tilde{y} = y - F(x_0) \approx F'(x_0)(x) + F''(x_0)(x,x) \tag{15}$$

which is a sum of a linear and a bilinear operator. In the first example, we have to solve a bilinear problem; in the second, we are going to replace equation (1) by its approximation (14) and solve the resulting problem. A third example from medical imaging is presented in Section 3. Finally, many parameter estimation problems for partial differential equations have a bilinear structure [4, 6, 14, 19]. We are now left with the problem of solving a 'linear–bilinear' equation of the type

$$F(f,\mu) = Af + B(f,\mu) = y \tag{16}$$

with two unknown functions f and μ. To be more general, we will allow μ to be different from f; this will be needed for the medical application in the last section. For convenience, we will denote the first argument of B always by f (i.e. $f_k, f_*, ...$) and the second by μ.

2 Convergence of the steepest descent method for minimizing the Tikhonov–Phillips functional for bilinear operator equations

2.1 Error estimates for Tikhonov–Phillips–Regularization

In this section we will collect some general results about the parameter choice for Tikhonov–Phillips regularization. We are looking for a $\bar{x}$-minimum-norm-solution of $F(x) = y$, i.e. a solution closest to $\bar{x}$. In the following, every minimizer of

$$\Phi_\alpha(x) := \|y^\delta - F(x)\|^2 + \alpha\|x - \bar{x}\|^2 , \tag{17}$$

is denoted by x_α^δ. We will restrict our attention to a priori parameter choice rules and give the following main result, which will be used further on:

Theorem 2.1 *Let F be a (nonlinear) continuous and weakly sequentially closed operator with convex definition area $D(F)$ and let x_* be a $\bar{x}$ minimum norm solution of $F(x) = y$. Moreover, let the following conditions hold:*

1. *F is Fréchet–differentiable*
2. *there exists a $\gamma \geq 0$ such that $\|F'(x_*) - F'(x)\| \leq \gamma\|x_* - x\|$ for all $x \in D(F)$ in a sufficiently large ball around x_**
3. *there exists $\omega \in Y$ satisfying $x_* - \bar{x} = F'(x_*)^*\omega$ and*
4. *$\gamma\|\omega\| < 1$.*

Then, for the choice $\alpha \sim \delta$, we obtain

$$\|x_\alpha^\delta - x_*\| = O(\sqrt{\delta}) \tag{18}$$

and

$$\|F(x_\alpha^\delta) - y^\delta\| = O(\delta) . \tag{19}$$

A proof can be found in [7], p.245. A somewhat closer inspection of the proof yields the estimates

$$\begin{aligned} \|x_\alpha^\delta - x_*\| &\leq \frac{\delta + \alpha\|\omega\|}{\sqrt{\alpha}\sqrt{(1-\gamma\|\omega\|)}} ; && (20) \\ \|F(x_\alpha^\delta) - y^\delta\| &\leq \delta + 2\alpha\|\omega\| . && (21) \end{aligned}$$

Setting $\alpha = c\delta$, (20) becomes

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

We are looking for a minimum of $f(c)$:

$$f'(c) = \frac{1}{\sqrt{(1-\gamma\|\omega\|)}} \frac{\sqrt{c}\|\omega\| - \frac{1}{2}(1+c\|\omega\|)c^{-1/2}}{c} = 0 \ ,$$

i.e.

$$c_{opt} = \frac{1}{\|\omega\|} \ . \tag{22}$$

In the following we will assume that c_{opt} is known explicitly and we will always choose $\alpha = c_{opt}\delta$. With this parameter choice, (20) and (21) become

$$\|x_\alpha^\delta - x_*\| \ \leq \ \frac{2\|\omega\|^{1/2}}{(1-\gamma\|\omega\|)^{1/2}}\delta^{1/2} \tag{23}$$

$$\|F(x_\alpha^\delta) - y^\delta\| \ \leq \ 3\delta \ . \tag{24}$$

We may note that $f(c_{opt}) \to 0$ for $\|\omega\| \to 0$.

2.2 A convexity property of the Tikhonov–Phillips–functional

Steepest descent methods are a useful tool to minimize a functional ϕ. They seldom fail to converge to at least a local minimum of ϕ. A property which ensures convergence to the global minimum is the convexity of the functional, cp. [16]. For a linear operator A, the Tikhonov–Phillips functional is a (global) convex functional. But this property will be lost as soon as the operator is nonlinear. As a consequence, Φ_α will have several local minima, and it is not clear where a minimization process will converge. In order to find a global minimizer of the functional

$$\Phi_\alpha(f,\mu) := \|y^\delta - Af - B(f,\mu)\|^2 + \alpha\|(f,\mu) - (\bar{f},\bar{\mu})\|^2 \tag{25}$$

with an a priori guess to a solution of (6), we want to use a steepest descent method in the following sections. Although (25) is not convex for all inputs (f,μ), we might show that it is at least a *locally* convex functional. To show this, we have to examine the second derivative of the function $\phi(t) = \Phi_\alpha((f,\mu) + t\cdot h)$, $t \in \mathbb{R}$, $h = (h_1,h_2)$. We get

$$\begin{aligned}
\phi_h(t) &= \|y^\delta - Af - t\cdot Ah_1 - B(f,\mu) - t\cdot(B(f,h_2) + B(h_1,\mu)) - t^2\cdot B(h_1,h_2)\|^2 \\
&\quad +\alpha\|(f,\mu) + t\cdot(h_1,h_2) - (\bar{f},\bar{\mu})\|^2 \\
&= \Phi_\alpha(f,\mu) \\
&\quad -2t\left(\langle y^\delta - Af - B(f,\mu), Ah_1 + B(h_1,\mu) + B(f,h_2)\rangle_Y + \alpha\langle (f,\mu) - (\bar{f},\bar{\mu}), (h_1,h_2)\rangle_{X_1\times X_2}\right) \\
&\quad +t^2\left(\|Ah_1 + B(h_1,\mu) + B(f,h_2)\|_Y^2 - 2\langle y^\delta - Af - B(f,\mu), B(h_1,h_2)\rangle_Y + \alpha\|(h_1,h_2)\|_{X_1\times X_2}^2\right) \\
&\quad +2t^3\langle Ah_1 + B(h_1,\mu) + B(f,h_2), B(h_1,h_2)\rangle_Y \\
&\quad +t^4\|B(h_1,h_2)\|^2 \ .
\end{aligned}$$

Especially for the Fréchet– derivative of the Tikhonov–Phillips functional we find $\Phi_\alpha{}'(f,\mu)h = \phi_h'(0)$. Defining the linear operators S_μ and T_f as

$$S_\mu(\cdot) \;:=\; A(\cdot) + B(\cdot,\mu) \tag{26}$$

$$T_f \;:=\; B(f,\cdot) \tag{27}$$

we get

$$\phi_h'(0) = -2\langle (S_\mu^*(y^\delta - Af - B(f,\mu)) - \alpha(f-\bar{f}), T_f^*(y^\delta - Af - B(f,\mu))) - \alpha(\mu-\bar{\mu}), (h_1,h_2)\rangle \;. \tag{28}$$

If we set

$$\varphi_h(t) = \Phi_\alpha((f_\alpha^\delta, \mu_\alpha^\delta) + th) \;, \tag{29}$$

where $(f_\alpha^\delta, \mu_\alpha^\delta)$ is a global minimizer of (25), we get $\varphi_h'(0) = 0$. The functional (25) is convex in a neighborhood of a global minimizer if $\varphi_h''(t) > 0$. We will actually show that $\varphi_h''(t) > 2\varepsilon =: 2\eta\delta$ holds under certain restrictions:

Theorem 2.2 *Let the conditions of Theorem 2.1 hold for an operator $F(x) = Af + B(f,\mu)$ with A, B as in (6). Moreover, let $(f_\alpha^\delta, \mu_\alpha^\delta)$ be a global minimizer of (25) and ω and η be chosen s.t.*

$$\|\omega\| < \min\left\{\frac{1}{\gamma}, \frac{1}{6+\eta}\right\} \;. \tag{30}$$

Then, with the parameter choice $\alpha = \frac{\delta}{\|\omega\|}$, there exists a ball $B_r(f_\alpha^\delta, \mu_\alpha^\delta)$ s.t. if $(f_\alpha^\delta, \mu_\alpha^\delta + th) \in B_r(f_\alpha^\delta, \mu_\alpha^\delta)$ for $t \in [0,t_0]$ and $\|h\| = 1$ then $\varphi_h''(t) > \eta\delta$. The radius r of B_r can be estimated by

$$r \geq r_1(\|\omega\|)\delta + r_2(\|\omega\|)\sqrt{\delta} \;, \tag{31}$$

where r_1, r_2 are defined in (46).

Proof:
It is sufficient to prove the existence of $t_{min} > 0$ such that $\varphi_h''(t) > 2\varepsilon$ holds for $t \in [0, t_{min}]$ and arbitrary h with $\|h\| = 1$.

First, we have

$$\begin{aligned}\varphi_h''(t) \;=\;& 2\left(\|Ah_1 + B(h_1,\mu_\alpha^\delta) + B(f_\alpha^\delta, h_2)\|_Y^2 - 2\langle y^\delta - Af_\alpha^\delta - B(f_\alpha^\delta,\mu_\alpha^\delta), B(h_1,h_2)\rangle_Y + \alpha\right)\\ &+12t^3\langle Ah_1 + B(h_1,\mu_\alpha^\delta) + B(f_\alpha^\delta,h_2), B(h_1,h_2)\rangle + 12t^4\|B(h_1,h_2)\|^2 \;.\end{aligned} \tag{32}$$

Using the abbreviations

$$\begin{aligned} y_1 &= y^\delta - Af_\alpha^\delta - B(f_\alpha^\delta,\mu_\alpha^\delta)\\ a &= Ah_1 + B(h_1,\mu_\alpha^\delta) + B(f_\alpha^\delta,h_2)\\ b &= B(h_1,h_2) \;,\end{aligned}$$

$u^2 = \langle u, u \rangle$, $|u| = \sqrt{u^2}$, $u \in \{a, b, y_1\}$, we get the shorter equation

$$\varphi_h^{''}(t) = 2(a^2 - 2\langle y_1, b \rangle + \alpha) + 12t\langle a, b \rangle + 12t^2 b^2 \ . \tag{33}$$

Because of (24), (12) we have

$$\begin{aligned} |\langle y_1, b \rangle| &\leq \|y^\delta - Af_\alpha^\delta - B(f_\alpha^\delta, \mu_\alpha^\delta)\| \|B\| \|(h_1, h_2)\| \\ &\leq 3\delta \end{aligned}$$

and therefore

$$\begin{aligned} \varphi_h^{''}(0) &\geq 2(\alpha - 2|\langle y_1, b \rangle|) \\ &\geq 2(\alpha - 6\delta) \\ &= 2\left(\frac{1}{\|\omega\|} - 6\right)\delta \ . \end{aligned} \tag{34}$$

We already know that $\varphi_h^{''}(0) \geq 0$ must hold, but (34) gives us an estimate of $\varphi_h^{''}(0)$, i.e. if $\|\omega\| < 1/6$ we do know a lower bound of $\varphi_h^{''}(0)$.
For directions h with $B(h_1, h_2) = 0$ we have

$$\varphi_h^{''}(t) = \varphi_h^{''}(0) \ ,$$

and for directions h with $Ah_1 + B(h_1, \mu_\alpha^\delta) + B(f_\alpha^\delta, h_2) = 0$

$$\varphi_h^{''}(t) = \varphi_h^{''}(0) + t^2 b^2 > \varphi_h^{''}(0) \ , t \in \mathbb{R}.$$

In order to show $\varphi_h^{''}(t) \geq 2\varepsilon > 0$ for all h with $\|h\| = 1$ and $t \in [0, t_{min}]$, we have to ensure $\varphi_h^{''}(0) - 2\varepsilon > 0$ first. As in (34), we get, by setting $\varepsilon = \eta\delta$,

$$\varphi_h^{''}(0) - 2\varepsilon \geq 2(\alpha - 2\langle y_1, b \rangle - \varepsilon) \tag{35}$$

$$\geq 2\left(\frac{1}{\|\omega\|} - 6 - \eta\right)\delta \geq 0 \tag{36}$$

as long as $\|\omega\| \leq 1/(6 + \eta)$. In practice, we first need to choose ω with

$$\|\omega\| < \min\left\{\frac{1}{\gamma}, \frac{1}{6}\right\} \ , \tag{37}$$

then choosing $\eta > 0$ s.t.

$$\|\omega\| < \min\{\frac{1}{\gamma}, \frac{1}{6 + \eta}\}$$

still holds. Then we are going to find $t_{min} > 0$ with $\varphi_h^{''}(t) \geq 2\eta\delta = 2\varepsilon$. The quadratic polynomial $p(t) := \varphi_h^{''}(t) - 2\varepsilon \to +\infty$ for $|t| \to \infty$. If t_1, t_2 denote the zeros of $p(t)$, then $p(t) \geq 0$ if $t_i \in \mathbb{C}$ or $t \notin (t_1, t_2)$. According to (36) the coefficient of t^0 is bigger than zero; thus both real zeros will have the same sign. Finally we observe that (cp. (32))

$$\varphi_h^{''}(-t) = \varphi_{-h}^{(2)}(t) \ ,$$

so it is sufficient to restrict our attention to the case of positive zeros

$$t_{1,2} = -\frac{\langle a,b\rangle}{2b^2} \begin{matrix}+\\-\end{matrix} \sqrt{\frac{|\langle a,b\rangle|^2}{4b^4} - \frac{a^2 - 2\langle y_1,b\rangle + \alpha - \varepsilon}{6b^2}}\ ,$$

i.e. the case $\langle a,b\rangle < 0$. Setting $\tau = -\frac{\langle a,b\rangle}{|a||b|}$, $\tau \in [0,1]$ and $t_{min} = \min\{t_1,t_2\}$, $t_{max} = max\{t_1,t_2\}$, we have

$$t_{min} = \tau\frac{|a|}{2|b|} - \sqrt{\frac{(3\tau^2-2)a^2 - 2(\alpha - \varepsilon - 2\langle y_1,b\rangle)}{12b^2}}\ . \tag{38}$$

Real zeros do only exist if

$$\frac{2}{3}\frac{a^2 + \alpha - 2\langle y_1,b\rangle - \varepsilon}{a^2} \le \tau \le 1 \tag{39}$$

holds. According to Vieta's Theorem,

$$t_{max}\cdot t_{min} = \frac{1}{6}\frac{a^2 + \alpha - 2\langle y_1,b\rangle - \varepsilon}{b^2} \tag{40}$$

holds. As a function of τ, t_{max} takes its maximum for $\tau = 1$ and can thus be estimated by

$$t_{max} \le \frac{|a|}{2|b|} + \sqrt{\frac{a^2 - 2(\alpha - \varepsilon - 2\langle y_1,b\rangle)}{12b^2}}$$

According to (35),(36), $\alpha - 2\langle y_1,b\rangle - \varepsilon$ is positive, and t_{max} can then be estimated by

$$t_{max} \le \frac{|a|}{2|b|} + \sqrt{\frac{a^2}{12b^2}} = \frac{|a|}{2|b|}\frac{\sqrt{3}+1}{\sqrt{3}}\ . \tag{41}$$

Then, again from (36), we get

$$\begin{aligned}\frac{a^2 + 2(c_{opt} - 6 - \eta)\delta}{6b^2} &\le \frac{a^2 + \alpha - 2\langle y_1,b\rangle - \varepsilon}{6b^2}\\ &= t_{min}\cdot t_{max}\\ &\le t_{min}\cdot\frac{|a|}{2|b|}\frac{\sqrt{3}+1}{\sqrt{3}}\end{aligned}$$

or

$$\begin{aligned} t_{min} &\ge 2(c_{opt} - 6 - \eta)\delta\frac{2\sqrt{3}|b|}{6(\sqrt{3}+1)|a|b^2} + \frac{2\sqrt{3}a^2|b|}{6(\sqrt{3}+1)b^2|a|}\\ &= \frac{2(c_{opt} - 6 - \eta)\delta}{\sqrt{3}(\sqrt{3}+1)|a||b|} + \frac{|a|}{\sqrt{3}(\sqrt{3}+1)|b|}\ . \end{aligned} \tag{42}$$

For a real zero t_{min}

$$0 \le (3\tau^2 - 2)a^2 - 2(\alpha - \varepsilon - 2\langle y_1,b\rangle)$$

holds, and especially for $\tau = 1$

$$\begin{aligned} 0 &\le a^2 - 2(\alpha - \varepsilon - 2\langle y_1,b\rangle)\\ &\le a^2 - 2(c_{opt} - 6 - \eta)\delta \end{aligned}$$

or

$$2\left(\frac{1}{\|\omega\|} - 6 - \eta\right)\delta \leq a^2 \ . \tag{43}$$

Now it follows ($|b| \leq 1$) that

$$\frac{|a|}{|b|} \geq \sqrt{2(\frac{1}{\|\omega\|} - 6 - \eta)}\sqrt{\delta} \ . \tag{44}$$

$|a|$ is bounded by

$$\begin{aligned} |a| &= \|Ah_1 + B(h_1, \mu_\alpha^\delta) + B(f_\alpha^\delta, h_2)\| \\ &\leq (\|A\| + \|B\|\|\mu_\alpha^\delta\|) + \|B\|\|f_\alpha^\delta\| \\ &\leq (\|A\| + 2\|B\|)\|(f_\alpha^\delta, \mu_\alpha^\delta)\| \end{aligned}$$

and so

$$\begin{aligned} \frac{1}{|a||b|} &\geq \frac{1}{|a|} \\ &\geq \frac{1}{\|A\| + 2\|B\|\|(f_\alpha^\delta, \mu_\alpha^\delta)\|} \ . \end{aligned} \tag{45}$$

Inserting (44), (45) in (42) we arrive at

$$\begin{aligned} t_{min} &\geq \frac{2(\frac{1}{\|\omega\|} - 6 - \eta)}{\sqrt{3}(\sqrt{3}+1)(\|A\| + 2\|B\|\|(f_\alpha^\delta, \mu_\alpha^\delta)\|)}\delta + \frac{\sqrt{2(\frac{1}{\|\omega\|} - 6 - \eta)}}{\sqrt{3}(\sqrt{3}+1)}\sqrt{\delta} \\ &=: r_1(\|\omega\|)\delta + r_2(\|\omega\|)\delta^{1/2} \ . \end{aligned} \tag{46}$$

□

From the definition of $r_2(\|\omega\|)$ we observe $r_2(\|\omega\|) \to \infty$ for $\|\omega\| \to 0$, which will be useful later on. We might add that the assumptions of Theorem 2.2 are only slightly more restrictive then in Theorem 2.1. Indeed, assumption 4 in Theorem 2.1 is already a condition to the smallness of $\|\omega\|$; and it only depends on the size of γ if (30) is a stronger restriction.

2.3 Analysis of steepest descent method

In this section it will be shown that the steepest descent algorithm will converge to a global minimum of the Tikhonov–Phillips functional, provided the starting value for the iteration process is in the area of convexity $B_r(f_\alpha^\delta, \mu_\alpha^\delta)$ and the step size of the iteration is chosen appropriately. Let $(f_k, \mu_k) \in B_r(f_\alpha^\delta, \mu_\alpha^\delta)$, r defined as in Theorem 2.2, and

$$h = (h_1, h_2) = (f_\alpha^\delta - f_k, \mu_\alpha^\delta - \mu_k) \ . \tag{47}$$

We define functions $\varphi_1(t), \varphi_2(t)$ by

$$\begin{aligned} \varphi_1(t) &= \Phi_\alpha((f_k,\mu_k)+th) && (48)\\ \varphi_2(t) &= \Phi_\alpha((f_\alpha^\delta,\mu_\alpha^\delta)-th)\ . && (49) \end{aligned}$$

According to (28) we have for an arbitrary $k=(k_1,k_2) \in X_1 \times X_2$

$$\begin{aligned} &\varphi_1'(0) = \Phi_\alpha{}'(f_k,\mu_k)k \\ &= -2\langle (S_{\mu_k}^*(y^\delta - Af_k - B(f_k,\mu_k)) - \alpha(f_k-\bar f), T_{f_k}^*(y^\delta - Af_k - B(f_k,\mu_k)) - \alpha(\mu_k-\bar\mu)), (k_1,k_2)\rangle\ . \end{aligned}$$

Setting

$$\nabla\Phi_\alpha(f_k,\mu_k) = (S_{\mu_k}^*(y^\delta - Af_k - B(f_k,\mu_k)) - \alpha(f_k-\bar f), T_{f_k}^*(y^\delta - Af_k - B(f_k,\mu_k)) - \alpha(\mu_k-\bar\mu))\ ,$$

the steepest descent method for minimizing Φ_α is then defined by

$$(f_{k+1},\mu_{k+1}) := (f_k,\mu_k) + \beta_k \nabla\Phi_\alpha(f_k,\mu_k)\ . \tag{50}$$

Here, β_k is a scaling parameter which has to be chosen appropriately in every iteration step.

Proposition 2.3 *Assume that $(f_k,\mu_k) \in B_r(f_\alpha^\delta,\mu_\alpha^\delta)$. Then an interval $I=(0,\beta_0)$ exists such that the iterate $(f_{k+1},\mu_{k+1}) = (f_k,\mu_k)+\beta_k\nabla\Phi_\alpha(f_k,\mu_k)$ is closer to $(f_\alpha^\delta,\mu_\alpha^\delta)$ as (f_k,μ_k) for $\beta_k \in I$:*

$$\|(f_\alpha^\delta,\mu_\alpha^\delta) - (f_{k+1},\mu_{k+1})\| < \|(f_\alpha^\delta,\mu_\alpha^\delta) - (f_k,\mu_k)\|\ .$$

Especially, $(f_{k+1},\mu_{k+1}) \in B_r(f_\alpha^\delta,\mu_\alpha^\delta)$.

Proof:
With the above definitions (47),(48),(49) we have

$$\begin{aligned} \varphi_1(t) &= \varphi_1(0) - 2\langle\nabla\Phi_\alpha(f_k,\mu_k),h\rangle t + a_h\cdot t^2 + b_h\cdot t^3 + c_h\cdot t^4 && (51)\\ \varphi_2(t) &= \varphi_2(0) + 2\langle\nabla\Phi_\alpha(f_\alpha^\delta,\mu_\alpha^\delta),h\rangle t + \bar a_h\cdot t^2 + \bar b_h\cdot t^3 + \bar c_h\cdot t^4 && (52) \end{aligned}$$

and some coefficients $a_h, b_h, c_h, \bar a_h, \bar b_h, \bar c_h$. Because $(f_k,\mu_k) \in B_r(f_\alpha^\delta,\mu_\alpha^\delta)$, $\varphi_2''(t) > 0$ holds for $0 \le t \le 1$. Moreover, $\varphi_2''(t) = \varphi_1''(1-t)$ and especially $2a_h = \varphi_1''(0) = \varphi_2''(1) > 0$. Thus, $\langle\nabla\Phi_\alpha(f_k,\mu_k),h\rangle > 0$:
If we assume $\langle\nabla\Phi_\alpha(f_k,\mu_k),h\rangle \le 0$, then $\varphi_1(0) < \varphi_1(t)$ would hold for small t. Now $\varphi_1(0) > \varphi_1(1)$, i.e. there $t_e \in [0,1]$ exists such that t_e is a local maximum of $\varphi_1(t)$. But, on the other hand, $\varphi_1''(t_e) > 0$, i.e. t_e is a local minimum, which is a contradiction. Now we have

$$\begin{aligned} \|(f_\alpha^\delta,\mu_\alpha^\delta) - (f_{k+1},\mu_{k+1})\|^2 &= \|(f_\alpha^\delta,\mu_\alpha^\delta) - (f_k,\mu_k)\|^2 + \beta^2\|\nabla\Phi_\alpha(f_k,\mu_k)\|^2 \\ &\quad -2\beta \underbrace{\langle\nabla\Phi_\alpha(f_k,\mu_k),(f_\alpha^\delta,\mu_\alpha^\delta)-(f_k,\mu_k)\rangle}_{=\langle\nabla\Phi_\alpha(f_k,\mu_k),h\rangle>0}\ . \end{aligned}$$

Thus

$$\|(f_\alpha^\delta,\mu_\alpha^\delta) - (f_{k+1},\mu_{k+1})\|^2 - \|(f_\alpha^\delta,\mu_\alpha^\delta) - (f_k,\mu_k)\|^2 < 0$$

if

$$g(\beta) := \beta^2 \|\nabla\Phi_\alpha(f_k, \mu_k)\|^2 - 2\beta\langle\nabla\Phi_\alpha(f_k, \mu_k), h\rangle < 0 \ ,$$

or

$$\beta < 2\frac{\langle\nabla\Phi_\alpha(f_k, \mu_k), h\rangle}{\|\nabla\Phi_\alpha(f_k, \mu_k)\|^2} \ .$$

The function $g(\beta)$ has its minimum at

$$\beta_{min} = \frac{\langle\nabla\Phi_\alpha(f_k, \mu_k), h\rangle}{\|\nabla\Phi_\alpha(f_k, \mu_k)\|^2} \ , \tag{53}$$

so we might choose $\beta_0 = \beta_{min}$.

□

Of course, usually β_0 is not available, so we have to find a lower bound.
For this, we need more information about $\varphi_1'(t)$:
We have already shown $\varphi_1'(0) < 0$. Because $\varphi_1(t)$ has a global minimum at $t = 1$, $\varphi_1'(1) = 0$ holds. Now assume $\varphi_1'(t) > 0$ holds for some $t \in [0, 1]$. Due to the continuity of $\varphi_1'(t)$, there must be $t_e \in [0, 1]$ with $\varphi_1'(t_e) > 0$ and $\varphi_1'(t_e) > \varphi_1'(t)$ for all $t \in [0, 1]$, i.e. t_e is a maximum. But this is in contradiction to $\varphi_1''(t) > 0$ for all $t \in [0, 1]$. Thus,

$$\varphi_1'(t) < 0 \qquad \text{for } t \in [0, 1) \ . \tag{54}$$

According to (51), $\varphi_1'(t)$ is computed by

$$\varphi_1'(t) = -2\langle\nabla\Phi_\alpha(f_k, \mu_k), h\rangle + 2a_h \cdot t + 3b_h \cdot t^2 + 4c_h \cdot t^3 \ ,$$

and $\langle\nabla\Phi_\alpha(f_k, \mu_k), h\rangle$ can then be estimated by

$$\begin{aligned} 2\langle\nabla\Phi_\alpha(f_k, \mu_k), h\rangle &= \underbrace{-\varphi_1'(t)}_{\geq 0} + 2\underbrace{a_h \cdot}_{\geq 2\varepsilon\|h\|^2} t + 3b_h \cdot t^2 + 4\underbrace{c_h \cdot}_{\geq 0} t^3 \\ &\geq 4\varepsilon\|h\|^2 t + 3b_h \cdot t^2 \ , \end{aligned}$$

($c_h = \|B(h_1, h_2)\|^2$), $|b_h|$ can be estimated by

$$\begin{aligned} |b_h| &= 2|\langle Ah_1 + B(h_1, \mu_k) + B(f_k, h_2), B(h_1, h_2)\rangle| \\ &\leq 2\|Ah_1 + B(h_1, \mu_k) + B(f_k, h_2)\|\|B(h_1, h_2)\| \\ &\leq 2((A + B\|\mu_k\|)\|h_1\| + B\|f_k\|\|h_2\|)B\|h_1\|\|h_2\| \\ &\leq 2(A + 2B\|(f_k, \mu_k)\|)\|h\|^3 \ , \end{aligned}$$

$$\begin{aligned} \|h\| &= \|(f_\alpha^\delta, \mu_\alpha^\delta) - (f_k, \mu_k)\| \\ &\leq \|(f_\alpha^\delta, \mu_\alpha^\delta) - (\bar{f}, \bar{\mu})\| + \|(f_k, \mu_k) - (\bar{f}, \bar{\mu})\| \\ &\leq \left(\frac{\|y^\delta - Af_k - B(f_k, \mu_k)\|^2}{\alpha} + \|(f_k, \mu_k) - (\bar{f}, \bar{\mu})\|^2\right)^{1/2} + \|(f_k, \mu_k) - (\bar{f}, \bar{\mu})\| =: \varrho \ , \end{aligned}$$

and this gives

$$|b_h| \leq 2(A + 2B\|(f_k, \mu_k)\|)\varrho\|h\|^2 =: k\|h\|^2 \ . \tag{55}$$

Finally we arrive at

$$2\langle \nabla\Phi_\alpha(f_k, \mu_k), h\rangle \geq \|h\|^2(4\varepsilon t - 3kt^2) = \|h\|^2 \cdot g(t)$$

The function $g(t)$ achieves its maximum t_{max} for

$$t_{max} = \frac{2\varepsilon}{3k} > 0 \ .$$

We have to consider two cases:

1. $t_{max} > 1$. Because of $g(0) = 0$ it follows $g(1) > 0$, and

$$2\langle \nabla\Phi_\alpha(f_k, \mu_k), h\rangle \geq \|h\|^2(4\varepsilon - 3k) > 0 \ .$$

2. $t_{max} < 1$, i.e.

$$2\langle \nabla\Phi_\alpha(f_k, \mu_k), h\rangle \geq \|h\|^2(4\varepsilon t_{max} - 3kt_{max}^2) > 0 \ .$$

By setting

$$c = \frac{1}{2}\min\{4\varepsilon - 3k, 4\varepsilon t_{max} - 3kt_{max}^2\} \tag{56}$$

we finally get

$$\langle \nabla\Phi_\alpha(f_k, \mu_k), h\rangle \geq c\|h\|^2 \ . \tag{57}$$

In a next step, $\|h\|$ will be estimated from below by $\varphi_1(0) - \varphi_1(1)$. According to (48), (49) we have $\varphi_1(t) = \varphi_2(1-t)$. Keeping in mind that $\nabla\Phi_\alpha(f_\alpha^\delta, \mu_\alpha^\delta) = 0$ holds, we get from (52)

$$\begin{aligned} \varphi_1(0) - \varphi_1(1) &= \varphi_2(1) - \varphi_2(0) \\ &= \bar{a}_h + \bar{b}_h + \bar{c}_h \ , \end{aligned}$$

with

$$\begin{aligned} \bar{a}_h &= \|Ah_1 + B(h_1, \mu_\alpha^\delta) + B(f_\alpha^\delta, h_2)\|^2 + \alpha\|h\|^2 - 2\langle y^\delta - Af_\alpha^\delta - B(f_\alpha^\delta, \mu_\alpha^\delta), B(h_1, h_2)\rangle \\ \bar{b}_h &= -2\langle Ah_1 + B(h_1, \mu_\alpha^\delta) + B(f_\alpha^\delta, h_2), B(h_1, h_2)\rangle \\ \bar{c}_h &= \|B(h_1, h_2)\| \end{aligned}$$

and h according to (47). Using the inequalities

$$\begin{aligned} \|(f_\alpha^\delta, \mu_\alpha^\delta) - (\bar{f}, \bar{\mu})\|^2 &\leq \frac{\|y^\delta - Af_k - B(f_k, \mu_k)\|^2}{\alpha} + \|(f_k, \mu_k) - (\bar{f}, \bar{\mu})\|^2 \ , \\ \|y^\delta - Af_\alpha^\delta - B(f_\alpha^\delta, \mu_\alpha^\delta)\| &\leq 3\delta \ , \\ \|Ah_1 + B(h_1, \mu_\alpha^\delta) + B(f_\alpha^\delta, h_2)\| &\leq (A + B\|\mu_\alpha^\delta\|)\|h_1\| + B\|f_\alpha^\delta\|\|h_2\| \\ &\leq (A + 2B\|(f_\alpha^\delta, \mu_\alpha^\delta)\|)\|h\| \\ &\leq (A + 2B(\|(f_\alpha^\delta, \mu_\alpha^\delta) - (\bar{f}, \bar{\mu})\| + \|(\bar{f}, \bar{\mu})\|))\|h\| \end{aligned}$$

$$\leq \left(A + 2B\left(\frac{\|y^\delta - Af_k - B(f_k,\mu_k)\|^2}{\alpha} + \|(f_k,\mu_k) - (\bar f,\bar\mu)\|^2\right)^{1/2}\right.$$

$$\left. + \|(\bar f,\bar\mu)\|\right)\|h\| =: \tilde c_1\|h\| \ ,$$

the quantities $\bar a_h, \bar b_h, \bar c_h$ can now be estimated as follows:

$$\begin{array}{rcl} \bar a_h & \leq & (\tilde c_1^2 + \alpha + 6B\delta)\|h\|^2 =: c_1\|h\|^2 \ , \\ |\bar b_h| & \leq & 2\tilde c_1 B\|h\|^3 =: c_2\|h\|^3 \\ \bar c_h & \leq & B\|h\|^4 \ . \end{array}$$

The minimal value of Φ_α, $\varphi_1(1)$, is usually not known, but

$$\varphi_{min} := \min_{t\in\mathbb{R}^+} \Phi_\alpha((f_k,\mu_k) + t\nabla\Phi_\alpha(f_k,\mu_k)) \tag{58}$$

is computable. Altogether this yields

$$\begin{array}{rcl} \varphi_1(0) - \varphi_{min} & \leq & \varphi_1(0) - \varphi_1(1) \\ & = & \varphi_2(1) - \varphi_2(0) \\ & \leq & c_1\|h\|^2 + c_2\|h\|^3 + B\|h\|^4 \ . \end{array} \tag{59}$$

Now, using (57), it follows for $\|h\| > 1$

$$\langle\nabla\Phi_\alpha(f_k,\mu_k), h\rangle \geq c \ , \tag{60}$$

and for $\|h\| \leq 1$ by using (59)

$$\langle\nabla\Phi_\alpha(f_k,\mu_k), h\rangle \geq \frac{c}{c_1 + c_2 + B}(\varphi_1(0) - \varphi_{min}) \ . \tag{61}$$

We might summarize our results in the following

Proposition 2.4 *Let (f_{k+1},μ_{k+1}) be computed by (50). If the scaling parameter β_k is chosen such that*

$$\beta_k \leq \min\left\{\frac{c}{\|\nabla\Phi_\alpha(f_k,\mu_k)\|^2}, \frac{c}{(c_1+c_2+B)}\frac{(\varphi_1(0)-\varphi_{min})}{\|\nabla\Phi_\alpha(f_k,\mu_k)\|^2}\right\} \tag{62}$$

holds, the new iterate (f_{k+1},μ_{k+1}) is closer to $(f_\alpha^\delta,\mu_\alpha^\delta)$ as (f_k,μ_k):

$$\|(f_{k+1},\mu_{k+1}) - (f_\alpha^\delta,\mu_\alpha^\delta)\| < \|(f_k,\mu_k) - (f_\alpha^\delta,\mu_\alpha^\delta)\| \ . \tag{63}$$

Proposition 2.5 *Let $\{(f_k,\mu_k)\}_{k\in\mathbb{N}}$ be the iterates of steepest descent for the Tikhonov–Phillips functional with β_k chosen according to (62) and $(f_0,\mu_0) \in B_r(f_\alpha^\delta,\mu_\alpha^\delta)$ with r being the radius of convexity of Φ_α. Then, there exists a constant M such that the the second derivative of*

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

is bounded:

$$|\varphi''(t)| \leq M\|\nabla\Phi_\alpha(f_k,\mu_k)\|^2 \ .$$

Proof:
According to the choice of β_k are all iterates in $B_r(f_\alpha^\delta, \mu_\alpha^\delta)$. Because $\|\nabla\Phi_\alpha(f_k,\mu_k)\|$ is bounded in $B_r(f_\alpha^\delta, \mu_\alpha^\delta)$ by a constant κ, the function $\varphi''(t)$ can be estimated as follows ($\nabla\Phi_\alpha(f_n,\mu_n) = (\bar{h}_1, \bar{h}_2)$):

$$\begin{aligned}
|\varphi''(t)| &= 2\Big|\|A\bar{h}_1 + B(\bar{h}_2,\mu_k) + B(f_k,\bar{h}_2)\|^2 + \langle y^\delta - Af_k - B(f_k,\mu_k), B(\bar{h}_1,\bar{h}_2)\rangle + \alpha\|\nabla\Phi_\alpha(f_k,\mu_k)\|^2 \\
&\quad +12t\langle A\bar{h}_1 + B(\bar{h}_2,\mu_k) + B(f_k,\bar{h}_2), B(\bar{h}_1,\bar{h}_2)\rangle + 12t^2\|B(\bar{h}_1,\bar{h}_2)\|^2\Big| \\
&\le \Big((\|A\| + 2\|B\|\|(f_k,\mu_k)\|)^2 + \|B\|\|y^\delta - Af_k - B(f_k,\mu_k)\| + \alpha\Big)\|\nabla\Phi_\alpha(f_k,\mu_k)\|^2 \\
&\quad +12t\|(\|A\| + 2\|B\|\|(f_k,\mu_k)\|)\|B\|\|\nabla\Phi_\alpha(f_k,\mu_k)\|^3 + 12t^2\|B\|^2\|\nabla\Phi_\alpha(f_k,\mu_k)\|^4 \\
&\le \Big((\|A\| + 2\|B\|\|(f_k,\mu_k)\|)^2 + \|B\|\|y^\delta - Af_k - B(f_k,\mu_k)\| + \alpha \\
&\quad +12\kappa\|(\|A\| + 2\|B\|\|(f_k,\mu_k)\|)\|B\|^2 + 12\kappa^2\|B\|^2\Big)\|\nabla\Phi_\alpha(f_k,\mu_k)\|^2 \\
&=: M_k\|\nabla\Phi_\alpha(f_k,\mu_k)\|^2 \ .
\end{aligned}$$

Obviously, M_k is bounded from above in $B_r(f_\alpha^\delta, \mu_\alpha^\delta)$, $M_k \le M$.

□

Now we can state our first main result.

Theorem 2.6 *Let $\{(f_k,\mu_k)\}_{k\in\mathbb{N}}$ be the iterates of steepest descent for the Tikhonov–Phillips functional,*

$$(f_{k+1},\mu_{k+1}) = (f_k,\mu_k) + \beta_k\nabla\Phi_\alpha(f_k,\mu_k) \ ,$$

$$\beta_k \le \min\left\{\frac{c}{\|\nabla\Phi_\alpha(f_k,\mu_k)\|^2}, \frac{c}{(c_1+c_2+B)}\frac{(\varphi_1(0)-\varphi_{min})}{\|\nabla\Phi_\alpha(f_k,\mu_k)\|^2}, \frac{1}{M}\right\} \tag{64}$$

and $(f_0,\mu_0) \in B_r(f_\alpha^\delta, \mu_\alpha^\delta)$ with r being the radius of convexity of Φ_α according to Theorem 2.2. Then (f_k,μ_k) converges to a global minimum of Φ_α:

$$(f_k,\mu_k) \to (f_\alpha^\delta, \mu_\alpha^\delta) \quad \textit{for } k \to \infty \ . \tag{65}$$

Proof:
If $(f_0,\mu_0) \in B_r(f_\alpha^\delta, \mu_\alpha^\delta)$ and the scaling parameters β_k are chosen according to (64), it follows from Proposition 2.4 that $(f_k,\mu_k) \in B_r(f_\alpha^\delta, \mu_\alpha^\delta)$. The sequence $\Phi_\alpha(f_k,\mu_k)$ is monotonously decreasing and bounded from below, thus there exists Φ_0 s.t. $\Phi_\alpha(f_k,\mu_k) \to \Phi_0$ for $k \to \infty$. We set for $t \in [0,1]$

$$\varphi(t) := \Phi_\alpha\left((f_n,\mu_n) + t\nabla\Phi_\alpha(f_n,\mu_n)\right)$$

and get

$$\varphi(t) - \varphi(0) = -t\|\nabla\Phi_\alpha(f_n,\mu_n)\|^2 + \frac{t^2}{2}\varphi''(\bar{t}) \ , \qquad 0 \le \bar{t} \le t \ . \tag{66}$$

According to Proposition 2.5, $|\varphi''(t)| \le M\|\nabla\Phi_\alpha(f_k,\mu_k)\|^2$ holds. Without loss of generality we can moreover assume $M > 1$.

Now suppose $\nabla\Phi_\alpha(f_n,\mu_n)$ does not converge to zero. Then there exists $\varepsilon_0 > 0$ s.t. for every $N \in \mathbb{N}$ exists $n > N$ with $\|\nabla\Phi_\alpha(f_n,\mu_n)\| \geq \varepsilon_0$. We get

$$\begin{aligned}\varphi(t)-\varphi(0) &\leq -t\|\nabla\Phi_\alpha(f_n,\mu_n)\|^2 + \frac{t^2}{2}M\|\nabla\Phi_\alpha(f_n,\mu_n)\|^2 \\ &= (-t+\frac{t^2}{2}M)\|\nabla\Phi_\alpha(f_n,\mu_n)\|^2 \ . \end{aligned} \tag{67}$$

If we set esp. $t = 1/M$, then

$$\varphi(1/M)-\varphi(0) \leq -\frac{1}{2M}\|\nabla\Phi_\alpha(f_n,\mu_n)\|^2 < 0 \ .$$

With $\varphi_{min} = \min\{\varphi(t) : t \in \mathbb{R}\}$ we find

$$\frac{1}{2M}\|\nabla\Phi_\alpha(f_n,\mu_n)\|^2 < \varphi(0)-\varphi(1/M) \leq \varphi(0)-\varphi_{min}$$

and

$$\begin{aligned}\frac{\varphi(0)-\varphi_{min}}{\|\nabla\Phi_\alpha(f_k,\mu_k)\|^2} &\geq \frac{1}{2M} \ , \\ \frac{1}{\|\nabla\Phi_\alpha(f_k,\mu_k)\|} &\geq \frac{1}{\kappa} \ , \end{aligned} \tag{68}$$

i.e. β_n is uniformly bounded from below, $\beta_n \geq \beta > 0$ (cp. 62). In addition to the assumption $\|\nabla\Phi_\alpha(f_k,\mu_k)\| \geq \varepsilon_0$ we can then assume that $|\varphi(0)-\Phi_0| \leq \frac{\beta}{4}\varepsilon_0^2 \leq \frac{\beta_n}{4}\varepsilon_0^2$ holds. Setting $t = \beta_n$ in (67) we get with $\beta_n \leq 1/M$

$$\varphi(\beta_n)-\varphi(0) \leq \beta_n(-1+\frac{\beta_n M}{2})\varepsilon_0^2 \tag{69}$$

$$\leq \beta_n(-1+\frac{1}{2})\varepsilon_0^2 = -\frac{\beta_n}{2}\varepsilon_0^2 \tag{70}$$

or

$$\varphi(\beta_n)-\Phi_0 \leq \varphi(0)-\Phi_0-\frac{\beta_n}{2}\varepsilon_0^2 \tag{71}$$

$$\leq \frac{\beta_n}{4}\varepsilon_0^2-\frac{\beta_n}{2}\varepsilon_0^2 = -\frac{\beta_n}{4}\varepsilon_0^2 < 0 \tag{72}$$

which is a contradiction to $\Phi_\alpha(f_n,\mu_n) \downarrow \Phi_0$.

Now, if we set $h = (f_\alpha^\delta - f_k, \mu_\alpha^\delta - \mu_k)$, we finally get by using (57)

$$c\|h\|^2 \leq \langle\nabla\Phi_\alpha(f_k,\mu_k),h\rangle \leq \|\nabla\Phi_\alpha(f_k,\mu_k)\|\|h\|$$

and

$$c\|h\| \leq \|\nabla\Phi_\alpha(f_k,\mu_k)\| \ , \tag{73}$$

with c defined in (56) and thus $(f_k,\mu_k) \to (f_\alpha^\delta,\mu_\alpha^\delta)$ for $k \to \infty$.

□

Remark 2.7

1. *Instead of computing a new scaling parameter β_n for every iteration, the lower bound β (cp. (68)) can be used as scaling parameter for all iteration steps.*

2. *Estimate (73) can be used to check the accuracy of the iterates during the iteration process.*

2.4 A global minimization process for Tikhonov–Phillips regularization

With a priori given regularization parameter α, the functional Φ_α is convex only in a neighborhood of the minimizing element $(f_\alpha^\delta, \mu_\alpha^\delta)$. It was shown in the last section that the steepest descent algorithm will converge to a global minimizer of Φ_α, provided a sufficient good starting value (f_0, μ_0) for the iteration is known. For arbitrary error level δ this will usually not be true: (f_0, μ_0) has to be in $B_r(f_\alpha^\delta, \mu_\alpha^\delta)$ and $r = r(\delta)$ with $r(\delta) \to 0$ for $\delta \to 0$. In practice, one might only know an approximation of the solution (f_*, μ_*) of (6). Therefore we propose the following ITP/SD algorithm (Iterated Tikhonov–Phillips/Steepest Descent):

- For given $\|\omega\|, \gamma, y^\delta, \delta$ and (f_0, μ_0) with $\|(f_0, \mu_0) - (f_*, \mu_*)\| \leq \varrho$ choose sufficiently big δ_0, α_0 and $q < 1$.
- Set $\alpha = \delta / \|\omega\|$.
- Choose n s.t. $q^{n+1}\alpha_0 < \alpha \leq q^n \alpha_0$.
- $(f_{\alpha_{-1}}^{\delta_{-1}}, \mu_{\alpha_{-1}}^{\delta_{-1}}) = (f_0, \mu_0)$.
- For $j = 0, .., n$
  - $\alpha_j = q^j \alpha_0$, $\delta_j = q^j \delta_0$.
  - $(f_0^j, \mu_0^j) = (f_{\alpha_{j-1}}^{\delta_{j-1}}, \mu_{\alpha_{j-1}}^{\delta_{j-1}})$.
  - Compute $(f_{\alpha_j}^{\delta_j}, \mu_{\alpha_j}^{\delta_j})$ as minimizing element of Φ_{α_j} with steepest descent and starting value (f_0^j, μ_0^j).
- Compute $(f_\alpha^\delta, \mu_\alpha^\delta)$ as minimizing element of Φ_α with steepest descent and starting value $(f_{\alpha_n}^{\delta_n}, \mu_{\alpha_n}^{\delta_n})$.

In the following we are going to describe how to choose the parameters δ_0, α_0 and q.
As in the previous sections, we will assume that δ, y^δ is given and the regularization parameter is defined by $\alpha = c_{opt}\delta$ with c_{opt} given in (22). Moreover, we will assume that an approximation (f_0, μ_0) to the solution (f_*, μ_*) of (6) is known,

$$\|(f_0, \mu_0) - (f_*, \mu_*)\| \leq \varrho \; . \tag{74}$$

According to (23), we have

$$\|(f_\alpha^\delta,\mu_\alpha^\delta)-(f_*,\mu_*)\| \le \frac{2\|\omega\|^{1/2}}{(1-\gamma\|\omega\|)^{1/2}}\delta^{1/2} =: d(\|\omega\|)\delta^{1/2} \ ,$$

and $d(\|\omega\|)\to 0$ for $\omega\to 0$. The functional $\varphi_h(t)=\Phi((f_\alpha^\delta,\mu_\alpha^\delta)+th)$ is convex as long as

$$\|th\| \le r_1(\|\omega\|)\delta + r_2(\|\omega\|)\delta^{1/2} \ ,$$

with $r_1(\|\omega\|), r_2(\|\omega\|)$ defined in (46). Obviously holds $r_i(\omega)\to\infty$ for $\|\omega\|\to 0$, $i=1,2$. Now, it follows for arbitrary $\tilde\delta>\delta$, $\tilde\alpha=c_{opt}\tilde\delta$

$$\|(\tilde f_\alpha^\delta,\tilde\mu_\alpha^\delta)-(f_0,\mu_0)\| - \|(f_*,\mu_*)-(\tilde f_\alpha^\delta,\tilde\mu_\alpha^\delta)\| \le \|(f_0,\mu_0)-(f_*,\mu_*)\| \le \varrho$$

or

$$\begin{aligned}\|(\tilde f_\alpha^\delta,\tilde\mu_\alpha^\delta)-(f_0,\mu_0)\| &\le \varrho + \|(f_*,\mu_*)-(\tilde f_\alpha^\delta,\tilde\mu_\alpha^\delta)\| \\ &\le \varrho + d(\|\omega\|)\tilde\delta^{1/2} \ .\end{aligned}$$

If we claim

$$\varrho + d(\|\omega\|)\tilde\delta^{1/2} \le r_2(\|\omega\|)\tilde\delta^{1/2} \ ,$$

then we arrive at the condition

$$\underbrace{\frac{\varrho}{r_2(\|\omega\|)-d(\|\omega\|)}}_{>0 \text{ for small } \|\omega\|} \le \tilde\delta^{1/2} \ . \tag{75}$$

If we choose

$$\delta_0^{1/2} := \frac{\varrho}{r_2(\|\omega\|)-d(\|\omega\|)} \qquad \alpha_0 = c_{opt}\delta_0 \ , \tag{76}$$

we get

$$\|(f_{\alpha_0}^{\delta_0},\mu_{\alpha_0}^{\delta_0})-(f_0,\mu_0)\| \le r_1(\|\omega\|)\delta_0 + r_2(\|\omega\|)\delta_0^{1/2} \ ,$$

i.e. (f_0,μ_0) lies in the convexity area of Φ_{α_0} and the method of steepest descent converges to $(f_{\alpha_0}^{\delta_0},\mu_{\alpha_0}^{\delta_0})$.
For $0<q<1$ we set $\alpha_1=q\alpha_0$, $\delta_1=q\delta_0$ and $(f_0^1,\mu_0^1)=(f_{\alpha_0}^{\delta_0},\mu_{\alpha_0}^{\delta_0})$. Then we get the estimate

$$\begin{aligned}\|(f_{\alpha_1}^{\delta_1},\mu_{\alpha_1}^{\delta_1})-(f_0^1,\mu_0^1)\| &\le \|(f_{\alpha_1}^{\delta_1},\mu_{\alpha_1}^{\delta_1})-(f_*,\mu_*)\| + \|(f_*,\mu_*)-(f_0^1,\mu_0^1)\| \\ &\le d(\|\omega\|)\delta_1^{1/2} + d(\|\omega\|)\delta_0^{1/2} && (77) \\ &= d(\|\omega\|)\left(\frac{1+\sqrt{q}}{\sqrt{q}}\right)\delta_1^{1/2} \ . && (78)\end{aligned}$$

Again, $\varphi_h(t)=\Phi_{\alpha_1}((f_{\alpha_1}^{\delta_1},\mu_{\alpha_1}^{\delta_1}+th)$ is a convex functional for

$$\|th\| \le r_1(\|\omega\|)\delta_1 + r_2(\|\omega\|)\delta_1^{1/2} \ .$$

For small $\|\omega\|$ (and independently of δ_1) holds

$$d(\|\omega\|)\left(\frac{1+\sqrt{q}}{\sqrt{q}}\right)\delta_1^{1/2} \le r_2(\|\omega\|)\delta_1^{1/2} , \tag{79}$$

and thus

$$\|(f_{\alpha_1}^{\delta_1}, \mu_{\alpha_1}^{\delta_1}) - (f_0^1, \mu_0^1)\| \le r_2(\|\omega\|)\delta_1^{1/2} , \tag{80}$$

(f_0^1, μ_0^1) lies in the area of convexity of Φ_{α_1} . The method of steepest descent converges to $(f_{\alpha_1}^{\delta_1}, \mu_{\alpha_1}^{\delta_1})$. Because all our above arguments did not depend on the size of δ_1 and α_1 but only on the size of $\|\omega\|$, this shows that our algorithm produces a sequence of functions $(f_{\alpha_j}^{\delta_j}, \mu_{\alpha_j}^{\delta_j})$, $j = 0..n$, which are always in the area of convexity of $\Phi_{\alpha_{j+1}}((f_{\alpha_{j+1}}^{\delta_{j+1}}, \mu_{\alpha_{j+1}}^{\delta_{j+1}}) + th)$. Finally, using $(f_{\alpha_n}^{\delta_n}, \mu_{\alpha_n}^{\delta_n})$ as starting value for the minimization of Φ_α, we find $(f_\alpha^\delta, \mu_\alpha^\delta)$.

Theorem 2.8 *If the parameters for the algorithm ITP/SD are chosen as described above and $\|\omega\|$ is small enough, then ITP/SD converges for every given y^δ, δ to a global minimizer of Φ_α.*

3 Numerical examples

We will give two numerical examples to demonstrate the necessity of our approach. The first one is a finite dimensional problem, where the minimizers can be computed analytically. The second problem comes from the medical imaging area.

3.1 A (simple) 2–dimensional problem

Let $F(x,y) : \mathbb{R}^2 \to \mathbb{R}$ be given by

$$F(x,y) = ax + bxy \qquad a, b \in \mathbb{R}\backslash\{0\}.$$

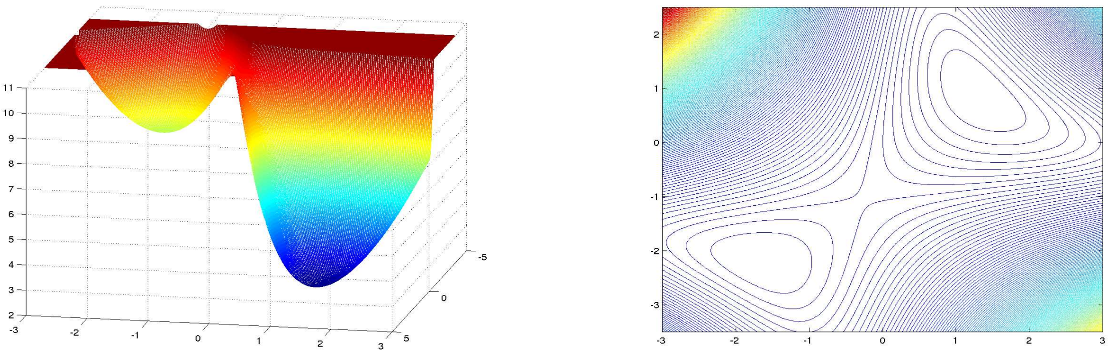
With given z, α, $\bar{x}$, $\bar{y}$ we might find the minimizer of the Tikhonov–Phillips functional

$$\Phi_\alpha(x,y) = (z - ax - bxy)^2 + \alpha((x-\bar{x})^2 + (y-\bar{y})^2). \tag{81}$$

For a minimum of Φ, the partial derivatives have to vanish:

$$\begin{aligned} 0 = \frac{\partial}{\partial x}\Phi(x,y) &= 2(z - ax - bxy)(-a - by) + 2\alpha(x - \bar{x}) \\ 0 = \frac{\partial}{\partial y}\Phi(x,y) &= -2(z - ax - bxy)bx + 2\alpha(y - \bar{y}) \end{aligned} \tag{82}$$

If we set $a = b = 1$, $z = 3$ and $\alpha = 0.8$, $(\bar{x}, \bar{y}) = (0,0)$, then we get three critical points as the



(a) (b)

Figure 1: (a) 3D-plot of $\Phi_{0.8}$. Clearly, the local minimum has a much higher value than the global minimum. (b) Contour plot of $\Phi_{0.8}$. Between the local and the global minimum lies the local maximum.

real solutions of (82):

$$\begin{aligned}(x_1, y_1) &= (-1.6118, -2.1876)\ ,\\ (x_2, y_2) &= (-1.0778, -0.2896)\ ,\\ (x_3, y_3) &= (+1.2969, +0.8899)\ .\end{aligned}$$

Figure 1 shows that (x_3, y_3) is the global minimum, $\Phi(x_3, y_3) = 2.2805$, and (x_2, y_2) is only a local minimum, $\Phi(x_2, y_2) = 7.0858$.
For a first numerical test the classical steepest descent method was used. As one might guess, the steepest descent iterates converged to one of the minima, but to which one was depending on the initial iterate (x_0, y_0). For example, with $(x_0, y_0) = (1, 1)$, the algorithm converged to the global minimum; but with $(x_0, y_0) = (-2, -1)$ only to the local one. In addition to this test, Newton's method for solving the system (82) was used, but here the iterates converged to all three critical points in dependence of the initial values. It is clear from our examples that these methods may fail for general operator equations, too.
For our ITP/SD algorithm and initial value $(x_0, y_0) = (-2, -1)$ the parameters $\alpha_0 = 4$, $q = 0.8$ were used. Table 1 shows that ITP/SD converges within a few steps to the global minimum of $\Phi_{0.8}$, even with a initial value where the classical steepest descent method or Newton's method converge only to a local minimum.
One might think that a good starting value (f_0, μ_0) of the steepest descent method for minimizing (81) might be $(f_0, \mu_0) = (\bar{f}, \bar{\mu})$. At least for our 2–dimensional example this seems to work in many cases . Choosing $(x_0, y_0) = (0, 0)$ as starting value for the above example, the steepest descent method converged to the global minimum. But we can give a counterexample: If we set $a = 0$, $b = 7$, $z = 3$, $\alpha = 0.1$, $(\bar{x}, \bar{y}) = (-2, 2)$ and $(x_0, y_0) = (\bar{x}, \bar{y})$, then the steepest descent method converges to $(x_1, y_1) = (-0.0095, 0.0095)$ with $\Phi_{0.1}(x_1, y_1) = 9.7962$ and our ITP/SD algorithm converges to $(x_2, y_2) = (-2.1944, -0.1944)$, $\Phi_{0.1}(x_1, y_1) = 0.4855$. It turns out that

Table 1: Results of ITP/SD after each outer iteration. j denotes the number of outer iterations, $\alpha_j = q^j \alpha_0$ and $(x_{\alpha_j}, y_{\alpha_j})$ is the minimizer of Φ_{α_j}.

j	α_j	x_{α_j}	y_{α_j}	$\Phi_{\alpha_j}(x_{\alpha_j}, y_{\alpha_j})$
0	4.0000	0.6971	0.3579	6.6726
1	2.8000	0.8925	0.5230	5.6882
2	1.9600	1.0552	0.6677	4.5944
3	1.3720	1.1765	0.7784	3.5543
4	0.9604	1.2631	0.8584	2.6659
5	0.8000	1.2969	0.8899	2.2805

(x_1, y_1) is only a critical point of (81), and (x_2, y_2) is a global minimum. Consequently, the results demonstrate that it is necessary to use our algorithm in order to find a global minimum of the Tikhonov–Phillips functional. Our example is only two–dimensional and it already has two minima. In case of an operator equation in infinite dimensions there might even exist many more extremal points.

3.2 An application from medical imaging

In the area of medical imaging, single–photon emission computed tomography (SPECT) plays an important role. Roughly speaking, it is used to find some abnormalities in the human body which can be indicated by the blood flow. The patient gets some radiopharmaceutical which will be transported by the blood and is supposed to enrich in a certain area of interest. After some time, the distribution of the radiopharmaceutical has to be recovered by measurements of the radioactivity outside the body. The connection between the measured data y and the distribution of the activity f is given by the Attenuated Radon Transform (ATRT):

$$y = R(f, \tilde{\mu})(s, \omega) = \int_{\mathbb{R}} f(s\omega^{\perp} + t\omega) e^{-\int_t^{\infty} \tilde{\mu}(s\omega^{\perp} + \tau\omega)\, d\tau} \, dt \ , \tag{83}$$

$s \in \mathbb{R}$, $\omega \in S^1$. As for the Radon Transform, the data are represented as line integrals over all possible unit vectors ω. The (usually also unknown) function $\tilde{\mu}$ 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. Attenuation correction has been an important area of nuclear medicine for more than two decades. Since the early 1990's, the move has been toward making some measurement of the attenuation distribution using a transmission source, and incorporating the reconstructed attenuation map, $\tilde{\mu}(x)$, in the reconstruction of the emission distribution. More recently, there has been renewed interest in methods that achieve good attenuation correction without the use of transmission measurements.

In [21], a bilinear approximation $\tilde{R}$ to R was introduced:

$$\tilde{R}(f, \mu) = \int_{\mathbb{R}} f(s\omega^{\perp} + t\omega) e^{-\int_t^{\infty} \mu_0(s\omega^{\perp} + \tau\omega)\, d\tau} (1 - \int_t^{\infty} \mu(s\omega^{\perp} + \tau\omega) d\tau) \, dt \ . \tag{84}$$

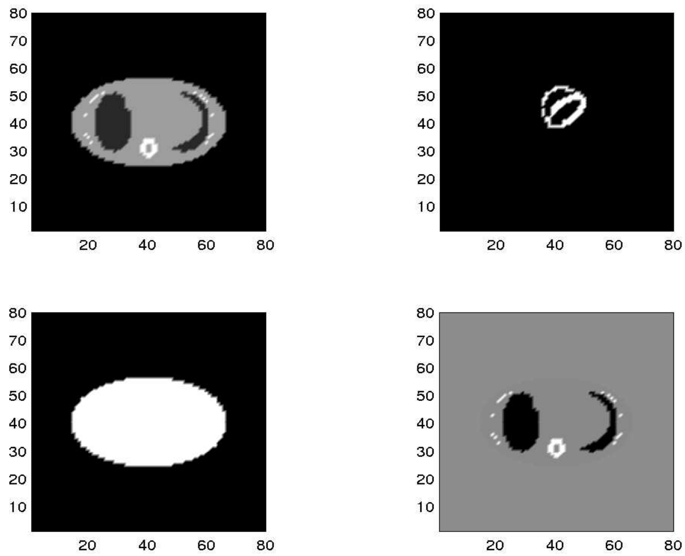


Figure 2: Attenuation distribution $\tilde{\mu}$ (top left), activity distribution f_* (top right), initial attenuation distribution μ_0 (bottom left), $\mu_* = \tilde{\mu} - \mu_0$ (bottom right).

Here, the exponential term in (83) was replaced by the first two terms of the Taylor–expansion for $\tilde{\mu} = \mu_0 + \mu$. For a fixed map μ_0, we are going to solve the equation

$$\tilde{R}(f, \mu) = y \tag{85}$$

instead of (83). This equation has the required form (16), so we might apply our ITP/SD algorithm.Test computations were done with the so called MCAT phantom [22]; the attenuation distribution is given by a model of the human body (with tissue, bones, lungs) and the activity f is assumed to have its support only in the heart (see Figure 2). The data $y = R(f, \tilde{\mu})$ was blurred with 10% Gaussian noise and then the steepest descent algorithm was used to minimize (25) with $(\bar{f}, \bar{\mu}) = (0, 0)$ for $\alpha = 7$. As starting value for the iteration $(f_0^1, \mu_0^1) = 0$ was taken. The reconstruction quality of $f_{rec} = f_\alpha^\delta$ $(\alpha = 7)$ looks good at a first glimpse, but μ_{rec} is far off μ. This is the well known phenomena that one gets a negative image of the activity function for the reconstructed attenuation map μ.
A first guess was that our algorithm converged to a local minimum of Φ_α instead to a global one close to the true solution $(f_*, \tilde{\mu}_*)$, but an examination of $\varphi(t) = \Phi_a((f_\alpha^\delta, \mu_\alpha^\delta) + t(f_* - f_\alpha^\delta, \mu_* - \mu_\alpha^\delta))$ (see Figure 4) shows that (f_*, μ_*) is actually close to a *local* minimum of Φ_α. We explain this behavior by the fact that we were actually looking for a solution close to zero $((\bar{f}, \bar{\mu}) = (0, 0))$. But

$$\|f_*\|^2 + \|\mu_*\|^2 \approx 9000 \gg \|f_\alpha^\delta\|^2 + \|\mu_\alpha^\delta\|^2 = 500\ ,$$

and this suggests that (f_*, μ_*) might be not the solution closest to zero. So we might state for SPECT that whenever an algorithm tries to reconstruct a solution closest to zero, one gets for the reconstruction of the attenuation function the above described phenomena of a negative

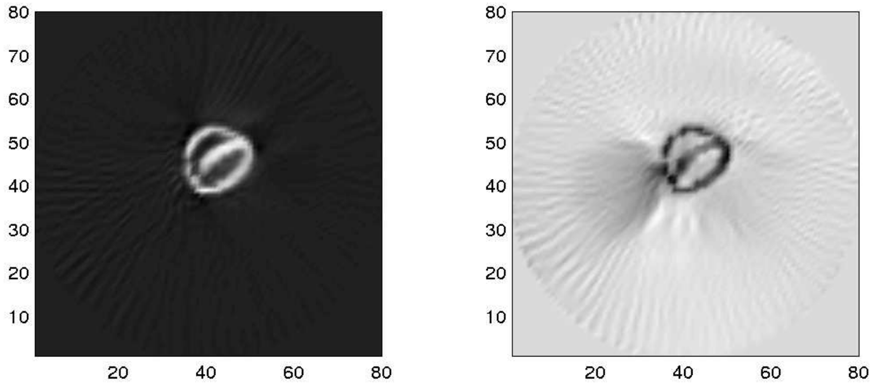


Figure 3: Reconstructed activity distribution f_{rec} (left), Reconstructed distribution μ (right), for $\alpha = 7$.

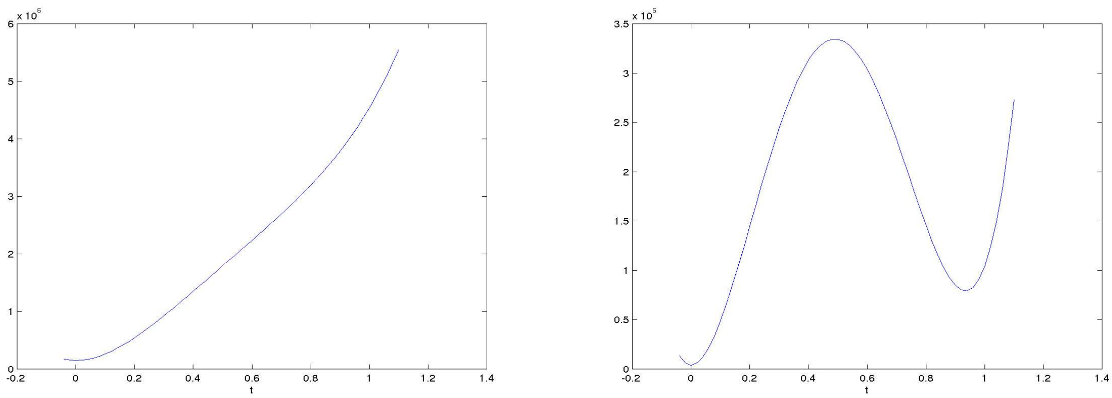


Figure 4: $\Phi_\alpha((f^\delta_\alpha, \mu^\delta_\alpha) + t(f_* - f^\delta_\alpha, \mu_* - \mu^\delta_\alpha))$ for $\alpha = 500$ (left) and $\alpha = 7$ (right). The solution (f_*, μ_*) is in the area of convexity of Φ_α only for the bigger α.

image of the activity function. In order to get a reconstruction for (f_*, μ_*), one has to choose $(\bar{f}, \bar{\mu})$ in (25) closer to the searched distributions.

In a second test we again tried to find $(f^\delta_\alpha, \mu^\delta_\alpha)$ with steepest descent for $\alpha = 7$, but now with a starting value (f_0^2, μ_0^2) which was close to the solution (f_*, μ_*). As the right picture in Figure 4 suggests, such a starting value for the iteration might not be in the area of convexity, and indeed the iteration converged only to a local minimum of Φ_α, and the according functional value of Φ_α was about six times as high as the value for the global minimum. Then, we used ITP/SD: At first, for $\alpha = 500$ and starting value (f_0^2, μ_0^2) , the minimizer $(f^\delta_\alpha, \mu^\delta_\alpha)$ was reconstructed and this value was taken as starting point for minimizing the Tikhonov–Phillips functional for $\alpha = 7$. In this case, the algorithm did again converge to the global minimum of Φ_α which was already reconstructed in our first attempt (cp. Table 2).

In a final test, we want to demonstrate the growth of the area of convexity with increasing α. To this end, we used the already computed minimizing functions of Φ_α for $\alpha \in \{7, 20, 500\}$ and computed the minimal real zero $t_{min}(h)$ of $\varphi_h^{(2)}(t) = {\Phi_\alpha}^{(2)}((f^\delta_\alpha, \mu^\delta_\alpha) + t(h_1, h_2))$ for a set of basis functions $h = (h_1, h_2) \in B$, $\|h\| = 1$. Then, the minimal zero $t_{min}(\alpha) = min\{t_{min}(h), h \in B\}$

Method	α	starting value	$\Phi_\alpha(f_\alpha^\delta, \mu_\alpha^\delta)$
SD	7	(f_0^1, μ_0^1)	3825
SD	7	(f_0^2, μ_0^2)	22956
ITP/SD	$\alpha_0 = 500$ $\alpha_1 = 7$	(f_0^2, μ_0^2)	3825

Table 2: Values of Φ_α for $\alpha = 7$, using steepest descent method (SD) with two different starting values for the iteration and ITP/SD ($(f_0^1, \mu_0^1) = (0,0)$, (f_0^2, μ_0^2) was chosen close to (f_*, μ_*)).

can be taken as the maximal radius of convexity of Φ_α. For our example, we found $t_{min}(\alpha = 7) = 1.5$, $t_{min}(\alpha = 20) = 1.7$ and $t_{min}(\alpha = 500) = 3.7$.

To summarize our numerical results, both examples have shown that it might be essential to use our algorithm in order to find a global minimum of the Tikhonov–Phillips functional. We have provided examples where other algorithms only found a local minimum and demonstrated the growth of the area of convexity with increasing regularization parameter.

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. Some newton type methods for the regularization of nonlinear ill–posed problems. *Inverse Problems*, (13):729–753, 1997.

[3] 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.

[4] 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.

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

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

[8] H.W. Engl. Discrepancy principles for Tikhonov regularization of ill posed problems leading to optimal convergence rates. *J. Optim.Theory Appl.*, (52):209–215, 1987.

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

[10] H.W. Engl, K. Kunisch, and A. Neubauer. Optimal a posteriori parameter choice for Tikhonov regularization for solving nonlinear ill-posed problems. *SIAM J. Numer. Anal.*, (30):1796–1883, 1993.

[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] 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.

[15] A. K. Louis. *Inverse und schlecht gestellte Probleme.* Teubner, Stuttgart, 1989.

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

[17] P. Maass, S. V. Pereverzev, R. Ramlau, and S. G. Solodky. An adaptive discretization scheme for Tikhonov–regularization with a posteriori parameter selection. *to appear in Numerische Mathematik*, 2000.

[18] P. Maaß and A. Rieder. Wavelet-accelerated tikhonov–regularisation with applications. In H.W. Engl, A.K. Louis, and W. Rundell, editors, *Inverse Problems in Medical Imaging and Nondestructive Testing*, pages 134–159. Springer, Wien, New York, 1997.

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

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

[21] 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.

[22] 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.

Berichte aus der Technomathematik ISSN 1435-7968

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

— Vertrieb durch den Autor —

Reports **Stand: 18. Januar 2001**

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.