SpezialForschungsBereich F 32

Karl–Franzens Universität Graz
Technische Universität Graz
Medizinische Universität Graz

On the Convergence of modified Landweber iteration for nonlinear inverse problems

H. Egger

SFB-Report No. 2010–017

April 2010

A–8010 GRAZ, HEINRICHSTRASSE 36, AUSTRIA

Supported by the
Austrian Science Fund (FWF)

SFB sponsors:

- **Austrian Science Fund** (FWF)
- **University of Graz**
- **Graz University of Technology**
- **Medical University of Graz**
- **Government of Styria**
- **City of Graz**

On the Convergence of modified Landweber iteration for nonlinear inverse problems

Herbert Egger

Abstract

This paper investigates the modified Landweber iteration for nonlinear inverse problems, and a projected variant of this method for problems with convex constraints. Optimal convergence rates are derived for different kinds of source conditions and under mild nonlinearity conditions on the operator. The theoretical results are illustrated by numerical results.

1 Introduction

In this paper, we consider the solution of nonlinear inverse problems of the form

$$F(x) = y^\delta,$$

where F is some nonlinear operator in Hilbert spaces, and y^δ denotes the possibly perturbed data. Such problems arise in many applications, e.g., in electric impedance tomography, in acoustic or electromagnetic scattering, and in more general parameter estimation problems.

In many cases, the inverse problem is ill-posed, in particular, the solution depends unstably on perturbations in the data, and we are especially interested in these cases. It is well-known, that for the solution of ill-posed problems, some regularization method has to be used [19, 2, 9].

From a practical point of view, iterative regularization methods [3, 16] are an attractive alternative to classical Tikhonov regularization, in particular for large-scale and nonlinear problems. The analysis of iterative regularization schemes for nonlinear problems depends on source conditions and additional conditions restricting the nonlinearity of the operator. While for smooth solutions, optimal convergence rates can be derived under simple Lipschitz conditions on the derivatives of the forward operator, the results for non-smooth solutions depend on stronger nonlinearity conditions.

An important aspect in the solution of inverse problems is the incorporation of a-priori conditions, e.g., via constraints. In this case, we would define the actual domain of F to be only a subset of the preimage space. Such constraints can be important, both, from an application point of view (e.g., to guarantee positivity of certain parameters), as well as from a mathematical point of view (for instance in order to to guarantee that the operator F is well-defined during the solution process). The solution of inverse problems with convex constraints has been investigated by several authors, e.g. in [17] for Tikhonov regularization, and in [3] for various iterative regularization methods; see also [15] for the analysis of a projected Gauß-Newton method, and [8] for results on the projected Landweber method.

In this paper, we consider the modified Landweber iteration [18] for the solution of possibly constrained nonlinear inverse problems. With our analysis, we intend to generalize the results obtained in [18] in several directions: First, we derive optimal convergence rates for smooth solutions under relaxed assumptions; in particular, we consider more general approximate source conditions utilized already in [3]; for a detailed study on such conditions see also [12, 11] and the references therein. If F additionally satisfies a tangential cone condition, we are able to derive optimal convergence rates also for Hölder type source conditions, similar to the results derived in [11] for Tikhonov regularization. To the best of our knowledge, these results are new for iterative regularization methods. In particular, the analysis of ordinary Landweber iteration [10] requires stronger range invariance conditions to guarantee similar convergence rates. The inductive nature of our estimates allows us to apply our analysis also to problems with convex constraints, which are solved by a projected variant of the modified Landweber iteration.

The outline of this manuscript is as follows: In Section 2, we state some basic assumptions on our problem, and we introduce the modified Landweber iteration. For the convenience of the reader, we also recall some of the convergence results of Landweber and the modified Landweber iteration. Section 3 is then devoted to the convergence analysis of the modified Landweber iteration for smooth solutions requiring only Lipschitz continuity of the derivative operator. The case of non-smooth solutions is investigated in Section 4, and optimal convergence rates are derived for Hölder type source conditions under the assumption that F only satisfies the tangential cone condition. In Section 5, the projected modified Landweber iteration is investigated, and convergence at optimal rates is established. For illustration of our results, we discuss in Section 6 a model problem from optical diffusion tomography, and we present results of numerical tests. The paper is concluded with a short summary.

2 Basic assumptions and preliminary results

In this paper, we consider the solution of nonlinear inverse problems of the form

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

where $F : D(F) \subset X \to Y$ is a continuously differentiable operator with domain $D(F) \subset X$. As usual, we assume that a solution $x^\dagger$ for the unperturbed data $y = F(x^\dagger)$ exists, and that a bound $\|y - y^\delta\| \le \delta$ on the data noise is available.

Remark 2.1. For notational convenience, we will assume throughout, that F is defined on the whole of X. However, our results can easily be localized, and it would suffice to consider only small balls around the solution. With the arguments used in [15, 6], one could even consider the case that the domain $D(F)$ has no interior points.

For the solution of the inverse problem (1), we consider the modified Landweber iteration [18]

$$x_{n+1} = x_n + F'(x_n)^*\big(y^\delta - F(x_n)\big) + \alpha_n(x_0 - x_n), \qquad n \ge 0, \tag{2}$$

where $x_0 \in D(F)$ is a given initial guess, and α_n denotes a sequence of non-negative parameters.

For later reference, let us recall a basic convergence result for this iteration. We assume that F' is Lipschitz continuous, i.e., that

$$\|F'(x_1) - F'(x_2)\| \le L\|x_1 - x_2\|. \tag{3}$$

This condition implies an estimate for the Taylor remainder, namely

$$\|F(x_1) - F(x_2) - F'(x_2)(x_1 - x_2)\| \le \frac{L}{2}\|x_1 - x_2\|^2.$$

For smooth solutions, the Lipschitz continuity of the derivative suffices to guarantee optimal convergence rates of the modified Landweber iteration.

Theorem 2.2 (cf. Theorem 3.1 in [18])**.** *Assume that $x^\dagger = x_0 + F'(x^\dagger)^* w$ for some $w \in Y$, and that F' is Lipschitz continuous with constant L. If $\|w\|$ and L are sufficiently small, and $\alpha_n = (n_0 + n)^{-\psi}$ for some $0 < \psi < 1$ and some n_0 large enough, than the iterates of the modified Landweber iteration satisfy $\|x_n - x^\dagger\| \le C\sqrt{\alpha_n}$ for some C independent of δ, as long as $\alpha_n \ge c\delta$ for some appropriate constant $c > 0$.*
If the iteration is stopped as soon as $\alpha_N \le c\delta$, then the optimal convergence ratse $\|x_N - x^\dagger\| = O(\sqrt{\delta})$ hold.

For ordinary Landweber iteration ($\alpha_n \equiv 0$), a similar convergence rates result could be proven only under the stronger "range invariance" condition

$$F'(x) = R_x F'(x^\dagger) \qquad \text{with} \qquad \|I - R_x\| < 1. \tag{4}$$

Similar nonlinearity conditions are frequently used for the analysis of iterative regularization methods for nonlinear inverse problems; see [16] for a detailed discussion.

Convergence of Landweber iteration (or the modified Landweber iteration) without rates can be proven under the weaker "tangential cone" condition [10]

$$\|F(x_1) - F(x_2) - F'(x_2)(x_1 - x_2)\| \leq \eta \|F(x_1) - F(x_2)\|, \qquad \eta < 1, \tag{5}$$

Note that the range invariance condition already implies the tangential cone condition.

For the Landweber iteration [18], or Newton-type iterations [14, 16] optimal convergence rates are also proven under Hölder type source conditions

$$x^\dagger = x_0 + (F'(x^\dagger)^* F'(x^\dagger))^\mu w, \qquad \text{for some } w \in X. \tag{6}$$

For the proof of these results, the range invariance condition (4) or similar strong nonlinearity conditions are required. To the best of our knowledge, a corresponding result for the modified Landweber iteration has not been proven so far.

3 Convergence rates for smooth solutions

In this section, we consider convergence of the modified Landweber iteration (2) for smooth solutions, under the assumption (3) that F has a Lipschitz continuous derivative . We first prove some a-priori estimates under "approximate source conditions" of the form

$$x^\dagger = T^* w + r, \qquad \text{for some } w \in Y \text{ and } r \in X. \tag{7}$$

Such source conditions have been introduced by Bakushinsky and co-workers; cf. [3]. As a direct consequence of these estimates, we obtain optimal rates provided that the "strong source condition"

$$x^\dagger = T^* w, \qquad \text{for some } w \in Y, \tag{8}$$

holds. Our results are derived under slightly simpler conditions as in [18], in particular, our results hold for more flexible choice of α_n.

3.1 Basic a-priori estimates

We start with deriving some basic identities, that will be used for the convergence proof in the following sections. The iteration error $x_n - x^\dagger$ of the modified Landweber iteration satisfies the following recursion

$$x_{n+1} - x^\dagger = (1 - \alpha_n)(x_n - x^\dagger) + F'(x_n)^*(y^\delta - F(x_n)) + \alpha_n(x_0 - x^\dagger). \tag{9}$$

Using this representation, we show now, that an estimate

$$\|x_n - x^\dagger\| \le C\sqrt{\alpha_n} \tag{10}$$

holds under appropriate assumptions on C and α_n. We will use an inductive argument; so let us assume that (10) already holds for iteration index n. To close the argument, let us consider the square of the iteration error:

$$\begin{aligned}\|x_{n+1} - x^\dagger\|^2 &= (1-\alpha_n)^2\|x_n - x^\dagger\|^2 + \|F'(x_n)^*(y^\delta - Tx_n)\|^2 \\ &\quad + \alpha_n^2\|x_0 - x^\dagger\|^2 + \underbrace{2(1-\alpha_n)(x_n - x^\dagger, F'(x_n)^*(y^\delta - F(x_n))}_{(a)} \\ &\quad + \underbrace{2(1-\alpha_n)\alpha_n(x_n - x^\dagger, x_0 - x^\dagger)}_{(b)} + \alpha_n(F'(x_n)^*(y^\delta - Tx_n, x_0 - x^\dagger).\end{aligned}$$

Next we derive estimates for the terms (a) and (b). Moving the adjoint operator to the left side of the scalar product, we obtain

$$\begin{aligned}(d) &= 2(1-\alpha_n)(F'(x_n)(x_n - x^\dagger), y^\delta - F(x_n)) \\ &\le 2(1-\alpha_n)\Big[-\|y^\delta - F(x_n)\|^2 + \big(\|R(x_n;x^\dagger)\| + \delta\big)\|y^\delta - F(x_n)\|\Big],\end{aligned} \tag{11}$$

where $R(x_n;x^\dagger) := F(x^\dagger) - F(x_n) - F'(x_n)(x^\dagger - x_n)$ denotes the Taylor remainder at x_n. In a similar manner, one sees that

$$\begin{aligned}(e) &= 2(1-\alpha_n)\alpha_n\big[(F'(x^\dagger)(x_n - x^\dagger), w) + (x_n - x^\dagger, r)\big] \\ &\le 2(1-\alpha_n)\alpha_n\Big[(\|y^\delta - F(x_n)\| + \|R(x^\dagger;x_n)\| + \delta)\|w\| + \|x_n - x^\dagger\|\|r\|\Big],\end{aligned} \tag{12}$$

where accordingly $R(x^\dagger;x_n) := F(x^\dagger) - F(x_n) - F'(x^\dagger)(x^\dagger - x_n)$. Using simple estimates for the remaining terms, collecting all parts that contain $\|y^\delta - Tx_n\|$, and assuming that $\alpha_n \le 1/4$, we obtain

$$\begin{aligned}\|x_{n+1} - x^\dagger\|^2 &\le (1-\alpha_n)^2\|x_n - x^\dagger\|^2 \\ &\quad - \frac{1}{2}\Big[\|y^\delta - F(x_n)\| - 2(\|R(x_n;x^\dagger)\| + \delta + \alpha_n\|w\|)\Big]^2 \\ &\quad + 2(\|R(x_n;x^\dagger)\| + \delta + \alpha_n\|w\|)^2 + 2\alpha_n(\|R(x^\dagger;x_n)\| + \delta)\|w\| \\ &\quad + 2\alpha_n|\|x_n - x^\dagger\|\|r\| + \alpha_n^2\|x_0 - x^\dagger\|^2.\end{aligned}$$

The quadratic term in brackets can be dropped. Using the Lipschitz estimate for the derivative to estimate the Taylor remainder terms, and applying the estimate $\|x_n - x^\dagger\| \le C\sqrt{\alpha}$, we further obtain

$$\begin{aligned}\|x_{n+1} - x^\dagger\|^2 &\le (1-\alpha_n)C^2\alpha_n + 6L^2C^4\alpha_n^2 + 6\delta^2 + 6\alpha_n^2\|w\|^2 \\ &\quad + 2\alpha_n(LC^2\alpha_n + \delta)\|w\| + 2\alpha_n C\sqrt{\alpha_n}\|r\| + \alpha_n^2\|x_0 - x^\dagger\|^2 \\ &\le C^2\alpha_n\Big[1 - \alpha_n\Big(2 - \alpha_n - 7L^2C^2 + \frac{7\delta^2}{C^2\alpha_n^2} + 9\frac{\|w\|^2}{C^2} + \frac{2\|r\|}{C\sqrt{\alpha_n}}\Big)\Big].\end{aligned} \tag{13}$$

If we assume that $\|w\|$ and L are sufficiently small, and α_n is sufficiently large compared to δ and $\|r\|^2$, then we can show that the term in parenthesis is bounded from below by one, and we obtain $\|x_{n+1} - x^\dagger\|^2 \leq C^2\alpha_n(1-\alpha_n)$, which can be used to define $\alpha_{n+1} \geq \alpha_n(1-\alpha_n)$.

Theorem 3.1. *Assume that $x^\dagger = x_0 + F'(x^\dagger)^*w + r$, and let $C \geq 7\|w\|$. Moreover, assume that F' is Lipschitz continuous, and $7LC \leq 1$.*
If $\alpha_0 \leq 1/4$ and $\alpha_{n+1} \leq \alpha_n(1-\alpha_n)$, then the iterates of the modified Landweber iteration satisfy $\|x_n - x^\dagger\| \leq C\sqrt{\alpha_n}$ as long as $\alpha_n \geq \frac{\delta}{\|w\|} + \frac{\|r\|^2}{\|w\|^2}$.

Proof. The result follows by induction from estimate (13), and the assumptions on the coefficients. □

Example 3.2. Let us consider the choice $\alpha_n = n^{-1}$. Then

$$\alpha_n(1-\alpha_n) \leq n^{-1}(1-n^{-1}) = \frac{n-1}{n^2} \leq \frac{n-1}{n^2-1} = (n+1)^{-1} = \alpha_{n+1}.$$

So this choice satisfies the requirements, apart from $\alpha_0 \leq 1/4$, which can easily be addressed by replacing $n \to n+4$. Any sequence of α_n's that decays slower than n^{-1} is possible as well. The restriction $\alpha_n \leq 1/4$ is actually not necessary. In fact, as long as $\alpha_n > 1/4$, the iterates stay in the vicinity of x_0, and this case can easily be treated with a different analysis.

Theorem 3.3. *Let the assumptions of Theorem 3.1 hold with $\|r\| = 0$, i.e., $x^\dagger = x_0 + F'(x^\dagger)^*w$. If the modified Landweber iteration is stopped as soon as $\alpha_N \leq \frac{\delta}{\|w\|}$, then $\|x_N - x^\dagger\| = O(\sqrt{\delta})$.*

Remark 3.4. Let us compare our result with the analysis of the modified Landweber iteration presented in [18]; see also Theorem 2.2. There, α_n was required to decay like $\alpha_n \sim n^{-\psi}$ for some $0 < \psi < 1$. This choice is also possible within our analysis. However, the parameters $\alpha_n \sim n^{-1}$ yield optimal convergence rates with fewer iterations, i.e., $N \sim \delta^{-1}$ in this case, which is also the optimal iteration count for Landweber iteration applied to linear problems.

Theorem 3.5. *Let the assumptions of Theorem 3.1 hold, and define N as the smallest index such that $\alpha_N \leq \frac{\delta}{\|w\|} + \frac{\|r\|^2}{\|w\|^2}$. Then $\|x_N - x^\dagger\| = O(\sqrt{\delta} + \|z\|)$.*

Remark 3.6. Note that the modified Landweber iteration with this choice of parameter is not a converging regularization method, unless $\|r\| = 0$ (or $\|r\| \to 0$ with $\delta \to 0$. The result however explains, why the iteration yields good results even for non-smooth solutions, as long as the noise level is not too small. Approximate source conditions of this kind have been used by [3] to analyze iterative regularization methods without requiring nonlinearity conditions on the operator. In the next section, we will use estimates of this kind to proof convergence and convergence rates also for Hölder type source conditions.

4 Hölder type source conditions

In the following, we consider convergence rates under Hölder type source conditions. We will proceed as follows: First, we derive estimate similar to that of Theorem 3.1 utilizing an additional nonlinearity condition. Then, we rephrase the Hölder type source condition as approximate source condition $x^\dagger = T^*w + r$, and derive bounds on $\|w\|$ and $\|r\|$. Similar relations for different kinds of source conditions were established in [12, 11]. Combining these two steps will imply the desired convergence results.

For the remainder of this section, we assume that F satisfies a tangential cone condition [10]

$$\|F(x_1) - F(x_2) - F'(x_2)(x_1 - x_2)\| \leq \eta \|F(x_1) - F(x_2)\|, \tag{14}$$

for all x_1, x_2 and some $\eta < 1$.

Remark 4.1. It suffices in principle, to require the condition in a ball around $x^\dagger$. For convenience, we again assume that the condition holds for all x_1, x_2. The following results could however be localized.

Using the tangential cone condition to estimate the norm of the Taylor remainder, we obtain for (11)

$$(d) \leq 2(1-\alpha)\Big(- \|y^\delta - F(x)\|^2(1-\eta) + (1+\eta)\delta\|y^\delta - F(x_n)\|\Big),$$

and (12) can be replaced by

$$(e) \leq 2(1-\alpha_n)\alpha_n\Big((1+\eta)(\|y^\delta - F(x_n)\| + \delta)\|w\| + \|x_n - x^\dagger\|\|r\|\Big).$$

Again, we collect all the terms containing $\|y^\delta - F(x_n)\|$, and assume that $\eta \leq 1/4$ and $\alpha_n \leq 1/8$ in the sequel. This yields

$$\begin{aligned}\|x_{n+1} - x^\dagger\|^2 \leq (1-\alpha_n)^2\|x_n - x^\dagger\|^2 - \frac{5}{16}\big[\|y^\delta - F(x_n)\| - 4(\delta - 2\alpha_n\|w\|)\big]^2 \\ + 5(\delta + \alpha_n\|w\|)^2 + \frac{5}{2}\alpha_n\delta\|w\| + 2\alpha_n\|x_n - x^\dagger\|\|r\| + \alpha_n^2\|x_0 - x^\dagger\|^2.\end{aligned}$$

Dropping the squared term, and inserting the estimates $\|x_n - x^\dagger\| \leq C\sqrt{\alpha_n}$, we now obtain

$$\|x_{n+1} - x^\dagger\|^2 \leq C^2\alpha_n\Big[1 - \alpha_n\big(2 - \alpha_n - \frac{12\delta^2}{C^2\alpha_n^2} - 13\frac{\|w\|^2}{C^2} - 2\frac{\|r\|}{C\sqrt{\alpha_n}}\Big],$$

which can be bounded by $C^2\alpha_n(1-\alpha_n)$, provided C and α_n are sufficiently large. Thus, the inductive step is completed if we choose $\alpha_{n+1} \geq \alpha_n(1-\alpha_n)$.

Theorem 4.2. *Let $x^\dagger = x_0 + F'(x^\dagger)^* w + r$, and F satisfy the tangential cone condition with $\eta \le 1/4$. Assume that $\alpha_0 \le 1/8$ and $\alpha_{n+1} \le \alpha_n(1-\alpha_n)$. Then for $C \ge 7\|w\|$, the iterates of the modified Landweber iteration* (2) *satisfy the a-priori estimate $\|x_n - x^\dagger\| \le C\sqrt{\alpha_n}$ as long as $\alpha_n \le \frac{\delta}{\|w\|} + \frac{\|r\|^2}{\|w\|^2}$.*

Proof. The result follows again by induction and the previous estimates. □

We intend to use this a-priori estimate, to derive convergence rates under Hölder type source conditions below. In order to be able to do that, let us relate the source condition $x^\dagger = (T^*T)^\mu z$ to the approximate source conditions considered above and in the previous section.

Lemma 4.3. *Let $T : X \to Y$ be a bounded linear operator, and assume that $x^\dagger = (T^*T)^\mu z$ for some $z \in X$ and some $0 < \mu \le 1/2$. Then for any $\theta > 0$ there exist elements $w \in Y$ and $r \in X$ with $\|w\| \le \theta^{2\mu-1}\|z\|$ and $\|r\| \le \theta^{2\mu}\|z\|$, such that $x^\dagger = T^*w + r$.*

Proof. For ease of presentation, let us assume that T is a compact operator, with singular value decomposition $\{\sigma_n, x_n, y_n\}$, and $x^\dagger = (T^*T)^\mu z$. Then

$$x^\dagger = \sum_{n\ge 1} \sigma_n^{2\mu}(z, x_n)x_n = \sum_{\sigma_n \ge \theta} \sigma_n \underbrace{\sigma_n^{2\mu-1}(z,x_n)}_{=:(w,w_n)} x_n + \sum_{\sigma_n<\theta} \underbrace{\sigma_n^{2\mu}(z,x_n)}_{=:(r,x_n)} x_n.$$

This defines two functions w and r which satisfy $x^\dagger = T^*w + r$. The estimates $\|w\| \le \theta^{2\mu-1}\|z\|$ and $\|r\| \le \theta^{2\mu}\|z\|$ follow directly. The case of non-compact operators can be considered with the same principle and spectral calculus. □

Remark 4.4. Similar statements are discussed in [12] and in [11] in a more general setting. In these works, the violation of the strong source condition is measured by the distance function $d(R) := \min_{w:\|w\|\le R} \|x^\dagger - T^*w\|$, and estimates for $d(R)$ are proven. The simple statement of Lemma 4.3 however suffices for our purposes.

Combining the previous results, we can now establish an a-priori estimate under Hölder type source conditions.

Theorem 4.5. *Let $x^\dagger = x_0 + \big(F'(x^\dagger)^*F'(x^\dagger)\big)^\mu z$, and assume that F satisfy the tangential cone condition* (14) *with $\eta \le 1/4$. If we choose $\alpha_0 \le 1/8$ and $\alpha_{n+1} \ge \alpha_n(1-\alpha_n)$, then $\|x_n - x^\dagger\| \le 7\delta^{\frac{2\mu-1}{2\mu+1}}\|z\|\sqrt{\alpha_n}$, as long as $\alpha_n \ge 2\|z\|^{-\frac{2}{2\mu+1}}\delta^{\frac{2}{2\mu+1}}$.*

Proof. Using Lemma 4.3 with $\theta = \delta^{\frac{1}{2\mu+1}}$, we obtain $x^\dagger = x_0 + F'(x^\dagger)^*w + r$ with $\|w\| \le \delta^{\frac{2\mu-1}{2\mu+1}}\|z\|$ and $\|r\| \le \delta^{\frac{2\mu}{2\mu+1}}\|z\|$. The result now follows by applying Theorem 4.2 with $C = 7\|w\| = 7\delta^{\frac{2\mu-1}{2\mu+1}}\|z\|$. □

Having established the a-priori estimate, we immediately obtain a convergence rates result by stopping the iteration appropriately.

Theorem 4.6. *Let the assumptions of Theorem 4.5 hold, and the stopping index N be the largest index such that $\alpha_N \le 2\|z\|^{-\frac{2}{2\mu+1}}\delta^{\frac{2}{2\mu+1}}$. Then the optimal convergence rates $\|x_N - x^\dagger\| = O\big(\delta^{\frac{2\mu}{2\mu+1}}\|z\|^{\frac{1}{2\mu+1}}\big)$ hold.*

Proof. The results follows directly from the estimates of Theorem 4.5. □

Remark 4.7. The parameter choice for the stopping index can be relaxed, i.e., it is not necessary to incorporate the norm of the source element in the stopping rule. This particular choice was made in the theorem to obtain also optimality with respect to the norm of the source element. The established rates coincide with the order optimal rates for Landweber iteration applied to linear problems; for details, see [9].

Remark 4.8. The choice $\alpha_n \sim n^{-1}$ is again possible, and using this choice, the number of iterations is bounded by $N = \delta^{-\frac{2}{2\mu+1}}$, which is the optimal iteration count obtained with Landweber iteration for linear problems.

Remark 4.9. The result of Theorem 4.6 is derived under weaker nonlinearity conditions than those required for the corresponding result for ordinary Landweber iteration [10]. There, a condition $F'(x) = R_x F'(x^\dagger)$ with $\|I - R_x\| < 1$ was used, which is stronger and implies the tangential cone condition. We do not believe, that the convergence rates could be proven for Landweber iteration when only requiring the tangential cone condition like we did here. In this respect, the modified Landweber iteration seems to be superior.

5 The projected modified Landweber method

The formulation of inverse problems often leads to problems with convex constraints. These constraints may come from physical reasoning, e.g. the positivity of parameters or fields, or it might be needed to ensure the forward operator to be well-defined; see, e.g., [15, 6].

In this section, we therefore consider inverse problems of the form: Find $x \in \mathcal{C} \subset X$ such that $F(x) = y^\delta$, where $\mathcal{C}$ denotes a closed, convex subset of X, and y^δ are the perturbed data, as before.

For the solution of this inverse problem, we again consider the modified Landweber iteration. In order to guarantee that the iterates satisfy the constraint (i.e. that x_n lies in the convex set $\mathcal{C}$ for all n), we use the following projected variant:

$$x_{n+1} = \mathcal{P}_{\mathcal{C}}\Big(\underbrace{x_n + F'(x_n)^*(y^\delta - F(x_n)) + \alpha_n(x_0 - x_n)}_{=:\tilde{x}_{n+1}}\Big), \qquad n \ge 0. \tag{15}$$

Here, $\mathcal{P}_{\mathcal{C}} : X \to \mathcal{C}$ denotes the metric projector onto $\mathcal{C}$. For our analysis, we utilize the non-expansiveness of the projection.

Lemma 5.1. *For any* x_1, $x_2 \in X$ *there holds* $\|\mathcal{P}_\mathcal{C}(x_1) - \mathcal{P}_\mathcal{C}(x_2)\| \le \|x_1 - x_2\|$.

Due to the inductive nature of our estimates, we can derive a-priori bounds for the projected modified Landweber method directly from the corresponding estimates for the method without projection.

Theorem 5.2. *Let* $x^\dagger = x_0 + F'(x^\dagger)^* w \in \mathcal{C}$, *and assume that* $x_0 \in \mathcal{C}$. *Then, under the conditions of Theorem 3.1, the iterates of the projected modified Landweber iteration* (15) *satisfy the a-priori estimate* $\|x_n - x^\dagger\| \le C\sqrt{\alpha_n}$, *provided that* $\alpha_n \ge \delta/\|w\|$.

Proof. Using the non-expansiveness of the projector, the definition of x_{n+1} and the fact that $x^\dagger \in \mathcal{C}$, we obtain

$$\|x_{n+1} - x^\dagger\| = \|\mathcal{P}_\mathcal{C}(\tilde{x}_{n+1}) - \mathcal{P}_\mathcal{C}(x^\dagger)\| \le \|\tilde{x}_{n+1} - x^\dagger\|.$$

The last term can be estimate as in Theorem 3.1, which completes the proof. □

Again, we obtain optimal convergence rates by stopping the iteration appropriately depending on the noise level.

Theorem 5.3. *Let the conditions of Theorem 5.2 hold. If the projected modified Landweber iteration* (15) *is stopped as soon as* $\alpha_N \le \delta/\|w\|$, *then the final iterate satisfies* $\|x_N - x^\dagger\| = O(\sqrt{\delta})$.

The corresponding result for Hölder-type source conditions follows similarly.

Theorem 5.4. *Let the assumptions of Theorem 4.6 be satisfied, and additionally* $x^\dagger \in \mathcal{C}$. *If the projected modified Landweber iteration is stopped as soon as* $\alpha_N \le 2\delta^{\frac{2}{2\mu+1}}$, *then the optimal rates* $\|x_n - x^\dagger\| = O(\delta^{\frac{2\mu}{2\mu+1}})$ *hold.*

Remark 5.5. A similar method of proof has been used in [3] to obtain optimal convergence rates for the projected iteratively regularized Gauß-Newton method using the approximate source condition $x^\dagger = x_0 + T^* w + r$; related results for operators only defined on $\mathcal{C}$ can be found in [15], see also [6] for a particular example. Let us note that we only require F to be well-defined on $\mathcal{C}$, therefore such generalizations are also possible here.

Remark 5.6. The projected Landweber iteration, which results from setting $\alpha_n \equiv 0$, has been considered in [8] for linear problems. Let us note, that in general, only weak convergence of the iterates to a stationary point can be proven. Under the assumption of attainability of the solution, and a weak source condition of Hölder type, we obtain here convergence of the full sequence for the modified Landweber iteration.

6 A test problem

In the following, we consider a standard inverse problem in optical diffusion tomography [1], namely the identification of absorption coefficients $c = c(x)$ from from boundary measurements of optical experiments.

6.1 Optical diffusion tomography

The propagation of light in highly scattering media is modelled by the elliptic boundary value problem

$$\begin{aligned} -\operatorname{div}(\kappa \nabla u) + cu &= 0 && \text{in } \Omega \\ \partial_n u + u &= q && \text{on } \partial\Omega, \end{aligned}$$

where u denotes the light intensity, the functions κ and c describe the optical properties of the material, and q models a light source at the boundary. We are interested in the identification of the absorption coefficient c, and therefore assume that the diffusion coefficient κ is known here.

Existence and uniqueness of a solution u to this partial differential equation is known for sources $q \in H^{-1/2}(\partial\Omega)$, provided that κ is uniformly positive and bounded, and that $c \in L^2(\Omega)$ is non-negative. Moreover, the solution can be shown to be uniformly bounded, i.e., $\|u\|_{H^1(\Omega)} \leq C\|q\|_{H^{-1/2}(\partial\Omega)}$ with a constant independent of the coefficient c.

In a typical experiment, several different light sources q are used to excite the object under investigation, and the resulting light intensities $u|_{\partial\Omega}$ at the boundary are measured by a set of detectors; hence, the outcome of such an experiment is a measurement (discretization) of the Robin-to-Dirichlet map, which we denote by M or $M(c)$ to emphasize the dependence on the parameter.

A mathematical formulation of such an aptical experiment can be given by a nonlinear *forward operator*

$$F : D(F) \subset L^2(\Omega) \to L(L^2(\partial\Omega), L^2(\partial\Omega)), \quad c \mapsto M(c),$$

which relates an absorption coefficient with the corresponding measurements. In order to guarantee solvability of the governing boundary value problem,but also for physical reasons, we define $D(F) := \mathcal{C} := \{c \in L^2(\Omega) : c \geq 0\}$, which is a closed, convex set in $L^2(\Omega)$.

The inverse problem considered in this section then consists of finding a function $c \in D(F)$, such that

$$F(c) = M^\delta, \tag{16}$$

where M^δ denotes the measured data. As usual, we assume that a parameter $c^\dagger$ corresponding to the unperturbed data $M = F(c^\dagger)$ exists, and that a bound $\|M^\delta - M\| \le \delta$ on the data noise is available.

Remark 6.1. The inverse problem (16) is known to have a unique solution; see [13] for details.

Let us summarize the main properties of the operator F.

Theorem 6.2 (see [6]). *The operator $F : D(F) \subset L^2(\Omega) \to L(L^2(\partial\Omega), L^2(\partial\Omega))$ is continuously differentiable, and the derivative satisfies a Lipschitz condition.*

Remark 6.3. Note that with respect to the L^2 topology, the set $D(F)$ has no interior points. One can generalize the analysis to this situation; cf. e.g. [15, 6]. For the problem considered here, one can also show that the operator F is in fact well-defined in a small neighborhood of $D(F)$, so the results derived in the previous sections can be applied in priciple without changes, if $\|x_0 - x^\dagger\|$ and $\|w\|$ in the source condition are sufficiently small.

We are thus in a situation, where the modified Landweber iteration provides optimal convergence rates for smooth solutions. Note, that we could not prove the tangential cone condition for the problem under investiagtion, so the results of Section 4 and 5 may not be applicable. However, this is the case for many applications, so we decided to use such a typical example for our numerical tests. Note

6.2 Discretization and numerical tests

For discretization of the governing boundary value problem, we consider a Galerkin method based on piecewise linear, continuous finite elements [5], and the space $L^2(\partial\Omega)$ of sources and measured intensities is discretized with piecewise linear, continuous functions on the boundary. The parameter function c is discretized with finite element functions as well.

Remark 6.4. This kind of discretization allows to derive exact derivatives and exact adjoints on the discrete level, so the results on convergence of the modified Landweber iteration can be applied also to the discrete setting, and guarantee convergence of the fully discrete methods. For details on the implementation, we refer to [7].

Let us now discuss the setup of our test problems: Se choose Ω to be the unit circle, and set $\kappa = 1$. We consider the reconstruction of two different parameters

$$\text{(ex1)} \qquad c(x,y) = \tfrac{1}{2}\big(\sin(\pi x)\cos(\pi y) + 1\big),$$
$$\text{(ex2)} \qquad c(x,y) = \max\big(\text{sign}[\sin(\pi x)\cos(\pi y) - 0.5], 0\big).$$

The first parameter is smooth and we expect to recover it with reasonable convergence rates. Although, we cannot show that this parameter satisfies a strong source condition, we can at least assume, that it satisfies an approximate source condition with $\|r\|$ not too large. For the second, non-smooth parameter, we expect to obtain less accurate reconstructions.

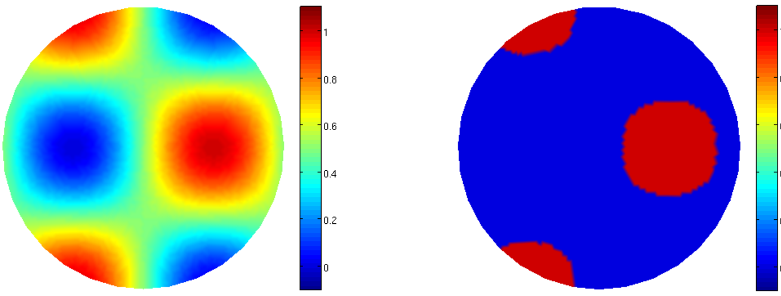

Figure 1: Coefficients $c^\dagger$ used for the numerical tests; left: $c^\dagger = \frac{1}{2}\big(\sin(\pi x)\cos(\pi y) + 1\big)$ used in test 1, and right: $c(x, y) = \max\big(\text{sign}[\sin(\pi x)\cos(\pi y) - 0.5], 0\big)$ for test 2.

For both test cases, the initial parameter was set to $c_0 \equiv \text{mean}(c^\dagger)$. The data $M = F(c^\dagger)$ were obtained by simulation, and additionally perturbed with random noise of size δ.

In Table 1, we list the reconstruction errors obtained with the modified Landweber iteration for test problem (ex1). Since we do not know the smoothness of our solution a-priori, we consider different choices for the stopping index N. Throughout, we use the choice $\alpha_n := n^{-1}$, which is admissible according to Example 3.2.

	$\alpha_N = \delta$		$\alpha_N = \delta^{4/3}$		$\alpha_N = \delta^{5/3}$		disc. princ.	
δ	α_N	$\frac{\|c_N - c^\dagger\|}{\|c^\dagger\|}$	α_N	$\frac{\|c_N - c^\dagger\|}{\|c^\dagger\|}$	α_N	$\frac{\|c_N - c^\dagger\|}{\|c^\dagger\|}$	$1/N$	$\frac{\|c_N - c^\dagger\|}{\|c^\dagger\|}$
0.32	0.2500	0.8760	0.2000	0.8506	0.1429	0.8203	0.5000	0.9125
0.16	0.1429	0.8196	0.0833	0.7689	0.0455	0.6996	0.0909	0.7731
0.08	0.0769	0.7551	0.0333	0.6713	0.0147	0.5930	0.0400	0.6840
0.04	0.0385	0.6850	0.0135	0.5818	0.0047	0.3800	0.0103	0.5421
0.02	0.0196	0.6230	0.0054	0.4107	0.0015	0.1912	0.0038	0.3184

Table 1: Convergence history of the modified Landweber iteration with smooth solution $c^\dagger = (\sin(\pi x)\cos(\pi y) + 1)/2$, different choices of the stopping index (left) and stopped according to the discrepancy principle with $\tau = 2.1$ (right).

Although, have not analyzed a-posteriori parameter choice strategies in this paper, we also compare with the results obtained by stopping with the discrepancy principle

$$N = \max\{n : \|F(x_n) - y^\delta\| \le \tau\delta\} \qquad \text{for some } \tau > 2.$$

The error after 50 iterations is approximately 0.62 for all test runs. Due to the different choices of the stopping index, the reconstruction errors are different however. Note, that for solutions satisfying only a weak source condition with index μ, one would set $\alpha_N = \delta^{\frac{2}{2\mu+1}} \sim \delta^2$, if μ is small. This explains why we obtain smaller errors with the choice $\alpha_N = \delta^{5/3}$ than for $\alpha_N = \delta$.

In the second test case, we consider the reconstruction of a discontinuou solution. Due to the Gibbs phenomenon, we expect, that projections are required to stay within the set of positive solutions, here. We compare the modified Landweber iteration, and its projected variatn with the standard Landweber method, which is again stopped according to the discrepancy principle with $\tau = 2.1$. For stopping the modified Landweber methods, we utilize again the a-priori rule $\alpha_N \sim \delta^{5/3}$, which corresponds to a smoothness index $\mu = 1/10$ in a Hölder type source condition.

	mod lw		proj mod lw		proj mod lw disc		lw disc	
δ	α_N	$\frac{\|c_N - c^\dagger\|}{\|c^\dagger\|}$	α_N	$\frac{\|c_N - c^\dagger\|}{\|c^\dagger\|}$	α_N	$\frac{\|c_N - c^\dagger\|}{\|c^\dagger\|}$	$1/N$	$\frac{\|c_N - c^\dagger\|}{\|c^\dagger\|}$
0.32	0.1429	0.9523	0.1429	0.9519	0.2500	0.9660	0.2500	0.9656
0.16	0.0455	0.9143	0.0455	0.8500	0.0455	0.8838	0.0333	0.8989
0.08	0.0147	0.8590	0.0147	0.8015	0.0149	0.7959	0.0099	0.8392
0.04	0.0047	0.7959	0.0047	0.7241	0.0014	0.6131	0.0017	0.7070
0.02	0.0015	0.6879	0.0015	0.6190	0.0005	0.4938	0.0007	0.5840
0.04	0.0100	0.8393	0.0100	0.7764	0.0100	0.7764	0.0100	0.8392

Table 2: Convergence history of the modified Landweber iteration, the projected modified Landweber iteration, and the standard Landweber method for the reconstruction of the discontinuus parameter $c^\dagger(x, y) = \max\big(\text{sign}[\sin(\pi x)\cos(\pi y) - 0.5], 0\big)$. The methods to the left are stopped according to $\alpha_N = \delta^{5/3}$, the two methods on the right are stopped with the discrepancy principle and $\tau = 2.1$. The last line displays the iteration errors after 100 iterations.

Since the solution is not smooth, and the problem under investgation is severely ill-posed, the convergence of the iteratio errors with respect to the noise level is very solw. Let us compare the errors obtained after 100 iterations for $\delta = 0.04$: The projected methods yields a smaller error, which is not surprising, since additional a-priori information was used, and due to the non-expansiveness of the projection, $\|\mathcal{P}_\mathcal{C}(x_n) - \mathcal{P}_C(x^\dagger)\| \leq \|x_n - x^\dagger\|$; see also the proof of Theorem 5.2. In the projected modified Landweber iteration, the projection step was performed in every iteration $n \geq 5$.

7 Discussion

In this paper, we investigated the modified Landweber iteration for the solution of nonlinear ill-posed problems. Additionally, we considered a projected variant for inverse problems with convex constraints.

For smooth solutions, we could show that the modified Landweber method converges with optimal rates provided that the operator under consideration has a Lipschitz continuous derivative. We derived a-priori estimates also under weaker, approximate source conditions, which allow a certain deviation of the solution (respectively the difference $x^\dagger - x_0$) from the range of $F'(x^\dagger)^*$. These results partly explain the good performance of the modified Landweber iteration also for non-smooth solutions and problems, for which stronger nonlinearity condition cannot be verified.

The convergence analysis was extended to Hölder type source conditions, where we used the relation of such conditions to the approximate source conditions. We derived optimal convergence rates requiring the operator F to satisfy only the tangential cone condition; similar results could be proven for the standard Landweber iteration only under stronger nonlinearity assumptions.

The inductive nature of our estimates allowed us to apply the convergence results immediately also to the projected modified Landweber iteration, and we obtained optimal convergence rates also for this method.

In this manuscript, we analyzed convergence only a-priori stopping rules. We think, that a-posteriori stopping rules based on Lepskij's principle can be applied as well; see [4] for a corresponding result for the Gauß-Newton method. From a numerical point of view, it would be desireable to use the discrepancy principle, since it requires considerably less iterations than the Lepskij principle. The analysis of efficient a-posteriori stopping rules is topic of future research.

References

[1] S. R. Arridge. Optical tomography in medical imaging. *Inverse Problems*, 15:R41–R93, 1999.

[2] A. B. Bakushinsky and A. V. Goncharskii. *Iterative methods for the solution of incorrect problems*. Nauka, Moscow, 1989.

[3] A. B. Bakushinsky and M. Yu Kokurin. *Iterative Methods for Approximate Solution of Inverse Problems*. Springer, Dordrecht, 2004.

[4] F. Bauer and T. Hohage. A Lepskij-type stopping rule for regularized Newton methods. *Inverse Problems*, 21:1975–1991, 2005.

[5] P. G. Ciarlet. *The Finite Element Method for Elliptic Problems*. North-Holland, Amsterdam, New York, Oxford, 1978.

[6] H. Egger and M. Schlottbom. Analysis and Regularization of Problems in Diffuse Optical Tomography. Technical Report, AICES-2009-23, RWTH Aachen University, 2009.

[7] H. Egger and M. Schlottbom. Towards reliable reconstructions in diffuse optical tomography. Technical Report, AICES-2010-8, RWTH Aachen University, 2010.

[8] B. Eicke. Iteration methods for convexly constrained ill-posed problems in Hilbert space. *Numer. Funct. Anal. Optim.*, 13:413–429, 1992.

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

[10] M. Hanke and A. Neubauer and O. Scherzer. A convergence analysis of Landweber iteration for nonlinear ill-posed problems. *Numer. Math.*, 72:21–35, 1995.

[11] T. Hein and B. Hofmann. Approximate source conditions for nonlinear ill-posed problems – chances and limitations. *Inverse Problems*, 25:035003, 2009.

[12] B. Hofmann. Approximate source conditions in Tikhonov–Phillips regularization and consequences for inverse problems with multiplication operators. *Math. Meth. Appl. Sci.*, 29:351–71, 2006.

[13] V. Isakov. *Inverse Problems for Partial Differential Equations, 2nd ed.*. Springer, 2006.

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

[15] B. Kaltenbacher and A. Neubauer. Convergence of projected iterative regularization methods for nonlinear problems with smooth solutions. *Inverse Problems*, 22:1105–1119, 2006.

[16] B. Kaltenbacher and A. Neubauer and O. Scherzer. *Iterative Regularization Methods for Nonlinear Problems.* de Guyter & Co, Berline, 2008.

[17] A. Neubauer. Tikhonov-regularization of ill-posed linear operator equations on closed convex sets. *J. Approx. Theory*, 53:304–320, 1988.

[18] O. Scherzer. A modified Landweber iteration for solving parameter estimation problems. *Appl. Math. Optim.*, 38:25–68, 1998.

[19] A. N. Tikhonov and V. Arsenin. *Solutions of Ill-Posed Problems.* Wiley, New York, 1977.