Taylor & Francis
Taylor & Francis Group

Check for updates

On a Family of Gradient-Type Projection Methods for Nonlinear Ill-Posed Problems

Antonio Leitão[a] and Benar F. Svaiter[b]

[a]Department of Mathematics, Federal University of Santa Catarina, Florianópolis, Brazil; [b]National Institute for Pure and Applied Mathematics, Estrada Dona Castorina, Rio de Janeiro, Brazil

ABSTRACT

We propose and analyze a family of successive projection methods whose step direction is the same as the Landweber method for solving nonlinear ill-posed problems that satisfy the *Tangential Cone Condition* (TCC). This family encompasses the Landweber method, the minimal error method, and the steepest descent method; thus, providing an unified framework for the analysis of these methods. Moreover, we define new methods in this family, which are convergent for the constant of the TCC in a range *twice as large* as the one required for the Landweber and other gradient type methods. The TCC is widely used in the analysis of iterative methods for solving nonlinear ill-posed problems. The key idea in this work is to use the TCC in order to construct special convex sets possessing a separation property, and to successively project onto these sets. Numerical experiments are presented for a nonlinear two-dimensional elliptic parameter identification problem, validating the efficiency of our method.

1. Introduction

In this article, we propose a family of successive orthogonal projection methods for obtaining stable approximate solutions to nonlinear ill-posed operator equations.

The *inverse problems*, we are interested in, consist of determining an unknown quantity $x \in X$ from the data set $y \in Y$, where X, Y are Hilbert spaces. The problem data y are obtained by indirect measurements of the parameter x, this process being described by the model $F(x) = y$, where $F : D \subset X \to Y$ is a non-linear ill-posed operator with domain $D = D(F)$.

In practical situations, the exact data y is not known. Instead, what is available is only approximately measured data $y^\delta \in Y$ satisfying

CONTACT Antonio Leitão ✉ acgleitao@gmail.com 🖳 Department of Mathematics, Federal University of Santa Catarina, P.O. Box 476, 88040-900, Florianópolis, Brazil; B. F. Svaiter ✉ benar@impa.br 🖳 National Institute for Pure and Applied Mathematics, Estrada Dona Castorina 110, 22460-320, Rio de Janeiro, Brazil.
Color versions of one or more of the figures in this article can be found online at www.tandfonline.com/lnfa.

$$||y^\delta - y|| \leq \delta, \tag{1}$$

where, $\delta > 0$ is the noise level. Thus, the abstract formulation of the inverse problems under consideration is to find $x \in D$ such that

$$F(x) = y^\delta. \tag{2}$$

The standard methods for obtaining stable solutions of the operator equation in (2) can be divided into two major groups, namely, *Iterative type* regularization methods [1–5] and *Tikhonov type* regularization methods [2, 6–10]. A classical and general condition commonly used in the convergence analysis of these methods is the *Tangent Cone Condition* (TCC) [3].

In this work, we use the TCC to define convex sets containing the local solutions of Equation (2) and devise a family of successive projection methods. The use of projection methods for solving linear ill-posed problems dates back to the 1970's (with the seminal works of Frank Natterer and Gabor Herman) [11–14]. The *combination* of Landweber iterations with projections onto a feasible set, for solving (2) with y^δ in a convex set was analyzed in [15] (see also [2] and the references therein).

The distinctive features of the family of methods proposed in this work are as follows:

- The basic method in this family outperformed, in our preliminary numerical experiments, the classical Landweber iteration [3] as well as the steepest descent iteration [16] (with respect to both the computational cost and the number of iterations);
- The family is generated by introducing relaxation in the stepsize of the basic method and such a family encompasses, as particular cases, the Landweber method, the steepest descent method, as well as the minimal error method [16]; thus, providing an unified framework for their convergence analysis;
- The basic method within the family converges for the constant in the TCC twice as large as required for the convergence of the Landweber and other gradient type methods.

In view of these features, the basic method within the proposed family is called the Projected Landweber (PLW) method. Although in the linear case the PLW method coincides with the minimal error method, in the nonlinear case these two methods are distinct.

The Landweber iteration was originally proposed for solving linear equations by using the method of successive approximations applied to the normal equations [5]. Its extension to non-linear equations was obtained by substituting the adjoint of the linear map by the Jacobian's adjoint of the

operator under consideration [3]. Such a method is named (nonlinear) Landweber, in the setting of ill-posed problems. Convergence of this method in the nonlinear case under the TCC was proven by Hanke et al. [3]. Convergence analysis for the steepest descent method and minimal error method (in the nonlinear case) can be found in [16].

Although Levenberg–Marquardt type methods are faster than gradient type methods, with respect to the number of iterations, gradient type methods have simpler and faster iteration formulas. Moreover, they fit nicely in Cimino and Kaczmarz type schemes. For these reasons, acceleration of gradient type methods is a relevant topic in the field of ill-posed problems.

The article is outlined as follows. In Section 2, we state the main assumptions and derive some auxiliary estimates required for the analysis of the proposed family of methods. In Section 3, we define the convex sets H_x (6), prove a special separation property (Lemma 3.1) and introduce our family of methods (8). Moreover, the first convergence analysis results are obtained, namely: monotonicity (Proposition 3.2) and strong convergence (Theorems 3.3 and 3.4) for the exact data case. In Section 4, we consider the noisy data case ($\delta > 0$). The convex sets H_x^δ are defined and another separation property is derived (Lemma 4.1). The discrepancy principle is used to define a stopping criteria (20), which is proved to be finite (Theorem 4.3). Monotonicity is proven (Proposition 4.2) as well as a stability result (Theorem 4.4) and a norm convergence result (Theorem 4.5). Section 5 is devoted to numerical experiments. In Section 6, we present final remarks and conclusions.

2. Main assumptions and preliminary results

In this section, we state our main assumptions and discuss some of their consequences, which are relevant for the forthcoming analysis. To simplify the notation, from now on we write

$$F_\delta(x) := F(x) - y^\delta \qquad \text{and} \qquad F_0(x) := F(x) - y. \tag{3}$$

Throughout this work we make the following assumptions, which are frequently used in the analysis of iterative regularization methods [2, 4, 7]:

A1 F is a continuous operator defined on $D(F) \subset X$, which has non-empty interior. Moreover, there exist constants C, $\rho > 0$ and $x_0 \in D(F)$ such that F', the Gateaux derivative of F, is defined on $B_\rho(x_0)$ and satisfies

$$\|F'(x)\| \leq C, \quad x \in B_\rho(x_0) \subset D(F) \tag{4}$$

(the point x_0 is be used as initial guess for our family of methods).

A2 The *local tangential cone condition* (TCC) [2, 4]

$$||F(\bar{x}) - F(x) - F'(x)(\bar{x} - x)||_Y \leq \eta ||F(\bar{x}) - F(x)||_Y, \quad \forall x, \bar{x} \in B_\rho(x_0)$$

(5)

holds for some $\eta < 1, x_0 \in X$, and $\rho > 0$.

A3 There exists an element $x^\star \in B_{\rho/2}(x_0)$ such that $F(x^\star) = y$, where $y \in Rg(F)$ are the exact data satisfying (1).

A4 The operator F is continuously Fréchet differentiable on $B_\rho(x_0)$.

Observe that in the TCC we require $\eta < 1$, instead of $\eta < 1/2$ as in classical convergence analysis for the nonlinear Landweber under this condition [2]. The TCC (5) represents a uniform assumption (on a ball of radius ρ) on the non-linearity of the operator F, and has interesting consequences (See [2, pp. 278–280] or [4, pp. 6 and Sec. 2.4 (pp. 26–29)]). Here, we discuss some of them.

Proposition 2.1. *If **A1** and **A2** hold, then for any $x, \bar{x} \in B_\rho(x_0)$*

1. $(1 - \eta)||F(x) - F(\bar{x})|| \leq ||F'(x)(x - \bar{x})|| \leq (1 + \eta)||F(x) - F(\bar{x})||$;
2. $\langle F'(x)^* F_0(x), x - \bar{x} \rangle \leq (1 + \eta)(||F_0(x)||^2 + ||F_0(x)||||F_0(\bar{x})||)$;
3. $\langle F'(x)^* F_0(x), x - \bar{x} \rangle \geq (1 - \eta)||F_0(x)||^2 - (1 + \eta)||F_0(x)||||F_0(\bar{x})||$.

If, additionally, $F_0(x) \neq 0$ then

$$(1 - \eta)||F_0(x)|| - (1 + \eta)||F_0(\bar{x})||$$
$$\leq ||F'(x)^*(x - \bar{x})|| \leq (1 + \eta)(||F_0(x)|| + ||F_0(\bar{x})||).$$

Proof. Item 1 follows immediately from the TCC and the triangle inequality, as proved in [2, Eq.(11.7)].

Direct algebraic manipulations yield

$$\langle F'(x)^* F(x), x - \bar{x} \rangle = \langle F(x), F'(x)(x - \bar{x}) \rangle$$
$$\leq (1 + \eta)||F(x)||||F(x) - F(\bar{x})||,$$

where, the inequality follows from Cauchy–Schwarz inequality and item 1. Likewise,

$$\langle F'(x)^* F_0(x), x - \bar{x} \rangle = \langle F_0(x), F'(x)(x - \bar{x}) \rangle$$
$$= \langle F_0(x), F_0(x) - F_0(\bar{x}) \rangle + \langle F_0(x), F_0(\bar{x}) - F_0(x) - F'(x)(\bar{x} - x) \rangle$$
$$\geq ||F_0(x)||^2 - ||F_0(x)||||F_0(\bar{x})|| - \eta ||F_0(x)||||F_0(x) - F_0(\bar{x})||,$$

where, the inequality follows from Cauchy–Schwarz inequality and the first inequality in this proof. Items 2 and 3 follow from the above inequalities and the inequality $||F_0(x) - F_0(\bar{x})|| \leq ||F_0(x)|| + ||F_0(\bar{x})||$. □

The next result relates to the solvability of operator equation $F(x) = y$ with exact data.

Proposition 2.2. *Let* **A1** *–* **A3** *be satisfied. For any* $x \in B_\rho(x_0)$, $F_0(x) = 0$ *if and only if* $F'(x)^* F_0(x) = 0$. *Moreover, for any* $(x_k) \in B_\rho(x_0)$ *converging to some* $\bar{x} \in B_\rho(x_0)$, *the following statements are equivalent:*

a) $\quad \lim\limits_{k \to \infty} ||F'(x_k)^* F_0(x_k)|| = 0;$

b) $\quad \lim\limits_{k \to \infty} ||F_0(x_k)|| = 0;$

c) $\quad F(\bar{x}) = y.$

Proof. See [4, pp. 279] for a proof of the first statement. For proving the second statement: the implication $(b) \Rightarrow (a)$ follows from **A1** and the hypothesis $(x_k) \in B_\rho(x_0)$; on the other hand, $(a) \Rightarrow (b)$ follows from Proposition 2.1, item 3 with $x = x_k$ and $\bar{x} = x^\star$; moreover, $(b) \Rightarrow (c)$ and $(b) \Leftarrow (c)$ follow from the hypothesis $\lim_{k \to \infty} ||x_k - \bar{x}|| = 0$ and **A1**. $\qquad \square$

Notice that the equivalence between (a) and (b) in Proposition 2.2 does not depend on the convergence of sequence (x_k). The next result provides a convenient way of rewriting the TCC (5) for $\bar{x} = x^\star \in F^{-1}(y)$ using notation (3).

Proposition 2.3. *Let* **A2** *be satisfied. If* $x^\star \in B_\rho(x_0) \cap F^{-1}(y)$ *then*

$$||y - y^\delta - F_\delta(x) - F'(x)(x^\star - x)|| \leq \eta ||y - y^\delta - F_\delta(x)||, \quad \forall x \in B_\rho(x_0).$$

3. A family of relaxed projection Landweber methods

In this section, we assume that exact data $y^\delta = y \in \mathrm{Rg}(F)$ are available, introduce a family of relaxed projection Landweber methods for the exact data case, and prove their convergence.

Define, for each $x \in D(F)$, the set

$$H_x := \{ z \in X \mid \langle z - x, F'(x)^* F_0(x) \rangle \leq -(1 - \eta) ||F_0(x)||^2 \}. \tag{6}$$

Note that H_x is either $\varnothing$, a closed half-space, or X. As we prove next, H_x has an interesting geometric feature: it contains all exact solutions of (2) in $B_\rho(x_0)$ and, whenever x is not a solution of (2), it does not contain x.

Lemma 3.1. (Separation): *Let* **A1** *and* **A2** *be satisfied. If* $x \in B_\rho(x_0)$ *then*

$$0 \geq (1 - \eta) ||F_0(x)||^2 + \langle F'(x)^* F_0(x), x^\star - x \rangle, \quad \forall x^\star \in B_\rho(x_0) \cap F^{-1}(y). \tag{7}$$

Consequently,
(1) $B_\rho(x_0) \cap F^{-1}(y) \subset H_x;$ $\qquad$ *(2)* $x \in H_x \iff F(x) = y.$

Proof. The first 2.1 item 3 with $\bar{x} = x^\star \in B_\rho(x_0) \cap F^{-1}(y)$. Items 1 and 2 are immediate consequences of the first statement and Definition (6). □

We are now ready to introduce our family of relaxed projection Landweber methods. Choose $x_0 \in X$ according to **A2** and **A3** and define, for $k \geq 0$, the sequence

$$x_{k+1} := x_k - \theta_k \lambda_k F'(x_k)^* F_0(x_k), \tag{8a}$$

$$\text{where } \theta_k \in (0,2), \lambda_k := \begin{cases} 0, & \text{if } F'(x_k)^* F_0(x_k) = 0 \\ \dfrac{(1-\eta)||F_0(x_k)||^2}{||F'(x_k)^* F_0(x_k)||^2}, & \text{otherwise.} \end{cases} \tag{8b}$$

In view of definition (6), the orthogonal projection of x_k onto H_{x_k} is $\hat{x} = x_k - \lambda_k F'(x_k)^* F_0(x_k)$ so that

$$x_{k+1} = x_k + \theta_k(\hat{x} - x_k) = (1-\theta_k)x_k + \theta_k\hat{x}.$$

We define the *PLW method* as (8) with $\theta_k = 1$ for all k. This choice amounts to taking x_{k+1} as the orthogonal projection of x_k onto H_{x_k}. The family of *relaxed projection Landweber methods* is obtained by choosing $\theta_k \in (0,2)$, which is equivalent to taking x_{k+1} as a relaxed orthogonal projection of x_k onto H_{x_k}.

Iteration (8) is well defined for all $x_k \in D(F)$; due to Proposition 2.2, this iteration becomes stationary at $x_{\tilde{k}} \in B_\rho(x_0)$, i.e. $x_k = x_{\tilde{k}}$ for $k \geq \tilde{k}$, if and only if $F(x_{\tilde{k}}) = y$.

In the next proposition an inequality is established, that guarantees the monotonicity of the iteration error for the family of relaxed projection Landweber methods in the case of exact data, i.e. $||x^\star - x_{k+1}|| \leq ||x^\star - x_k||$, whenever $\theta_k \in (0,2)$.

Proposition 3.2. *Let **A1** – **A3** hold true. If $x_k \in B_\rho(x_0), F'(x_k)^* F(x_k) \neq 0$, and θ_k and x_{k+1} are as in (8), then*

$$||x^\star - x_k||^2 \geq ||x^\star - x_{k+1}||^2 + \theta_k(2-\theta_k)\left((1-\eta)\frac{||F_0(x_k)||^2}{||F\prime(x_k)^* F_0(x_k)||}\right)^2,$$

for all

$$x^\star \in B_\rho(x_0) \cap F^{-1}(y.)$$

Proof. If $x_k \in B_\rho(x_0)$ and $F'(x_k)^* F(x_k) \neq 0$ then x_{k+1} is a relaxed orthogonal projection of x_k onto the H_{x_k} with a relaxation factor θ_k. The conclusion follows from this fact, Lemma 3.1, iteration formula (8) and the properties of relaxed metric projections onto arbitrary sets (see, e.g., [17, Lemma 3.13, pp. 21–22].) □

Direct inspection of the inequality in Proposition 3.2 shows that the choice $\theta_k \in (0, 2)$, as prescribed in (8), guarantees decrease of the iteration error $||x^\star - x_k||$, while $\theta_k = 1$ yields the greatest *estimated* decrease on the iteration error.

We are now ready to state and prove the main results of this section: Theorem 3.3 gives a sufficient condition for strong convergence of the family of relaxed projection Landweber methods (for exact data) to *some* point $\bar{x} \in B_\rho(x_0)$. Theorem 3.4 gives a sufficient condition for strong convergence of this family of methods to *a solution* of $F(x) = y$, shows that steepest descent, minimal error, as well as Landweber method are particular instances of methods belonging to this family, and proves convergence of these three methods within this framework.

Recall that the steepest descent method (SD) is given by

$$x_{x+1} = x_k - \frac{||F'(x_k)^* F_0(x_k)||^2}{||F'(x_k)F'(x_k)^* F_0(x_k)||^2} F'(x_k)^* F_0(x_k),$$

while the minimal error method (ME) is given by

$$x_{x+1} = x_k - \frac{||F_0(x_k)||^2}{||F'(x_k)^* F_0(x_k)||^2} F'(x_k)^* F_0(x_k).$$

Theorem 3.3. *If* **A1**–**A3** *hold true, then the sequences* $(x_k), (\theta_k)$ *as specified in* (8) *are well defined and*

$$x_k \in B_{\rho/2}(x^\star) \subset B_\rho(x_0), \forall k \in \mathbb{N}. \tag{9}$$

If, additionally, $\sup \theta_k < 2$, *then*

$$(1-\eta)^2 \sum_{k=0}^{\infty} \theta_k \frac{||F_0(x_k)||^4}{||F'(x_k)^* F_0(x_k)||^2} < \infty \tag{10}$$

and (x_k) *converges strongly to some* $\bar{x} \in B_\rho(x_0)$.

Theorem 3.4. *Let* **A1**–**A3** *hold true, and the sequences* $(x_k), (\theta_k)$ *be defined as in* (8). *The following statements hold:*

a. *If* $\inf \theta_k > 0$ *and* $\sup \theta_k < 2$, *then* (x_k) *converges to some* $\bar{x} \in B_\rho(x_0)$ *solving* $F(\bar{x}) = y$.

b. *If* **A2** *holds with* $\eta < 1/2$ *and*

$$\theta_k := (1-\eta)^{-1} \frac{||F'(x_k)^* F_0(x_k)||^2}{||F_0(x_k)||^2} \cdot \frac{||F'(x_k)^* F_0(x_k)||^2}{||F'(x_k)F'(x_k)^* F_0(x_k)||^2},$$

then $0 < \theta_k \leq (1-\eta)^{-1} < 2$, *iteration* (8) *reduces to the steepest descent method and* (x_k) *converges to some* $\bar{x} \in B_\rho(x_0)$ *solving* $F(\bar{x}) = y$.

a. If **A1** and **A2** hold with $C \leq 1$ and $\eta < 1/2$, respectively, and

$$\theta_k := (1-\eta)^{-1} \frac{||F'(x_k)^* F_0(x_k)||^2}{||F_0(x_k)||^2},$$

then $0 < \theta_k \leq (1-\eta)^{-1} < 2$, iteration (8) reduces to the nonlinear Landweber iteration and (x_k) converges to some $\bar{x} \in B_\rho(x_0)$ solving $F(\bar{x}) = y$.

a. If **A1** and **A2** hold with $C \leq 1$ and $\eta < 1/2$, respectively, and $\theta_k := (1-\eta)^{-1}$, then iteration (8) reduces to the nonlinear minimal error method and (x_k) converges to some $\bar{x} \in B_\rho(x_0)$ solving $F(\bar{x}) = y$.

Proof. (Theorem 3.3) Assumption **A3** guarantees the existence of $x^\star \in B_{\rho/2}(x_0)$, a solution of $F(x) = y$. It follows from **A3** that (9) holds for $k = 0$. Suppose that the sequence (x_k), is well defined up to k_0 and that (9) holds for $k = k_0$. It follows from **A1** that $x_{k_0} \in D(F)$, so that x_{k_0+1} is well defined while it follows from (8) and Proposition 3.2 that (9) also holds for $k = k_0 + 1$.

To prove the second part of the theorem, suppose that $b := \sup \theta_k < 2$. At this point, we have to consider two separate cases:

Case I: $F(x_{\tilde{k}}) = y$ for some $\tilde{k} \in \mathbb{N}$.

It follows from (9), Proposition 2.2 and (8), that $x_j = x_{\tilde{k}}$ for $j \geq \tilde{k}$, and we have trivially strong convergence of (x_k) to $\bar{x} = x_{\tilde{k}}$ (which, in this case, is a solution of $F(x) = y$).

Case II: $F(x_k) \neq y$, for all k.

It follows from (9) and Proposition 2.2 that $F'(x_k)^* F_0(x_k) \neq 0$ for all k. According to (8b)

$$\lambda_k := (1-\eta)||F_0(x_k)||^2 ||F'(x_k)^* F_0(x_k)||^{-2}. \tag{11}$$

Since $0 < \theta_k \leq b < 2$ for all k, $(2-\theta_k)\theta_k \geq (2-b)\theta_k > 0$, for all k. Therefore, it follows from Proposition 3.2 that

$$||x^\star - x_k||^2 + (2-b)\theta_k (1-\eta)^2 \sum_{j=0}^{k-1} \left(\frac{||F_0(x_j)||^2}{||F'(x_j)^* F_0(x_j)||} \right)^2 \leq ||x^\star - x_0||^2,$$

for all $x^\star \in B_\rho(x_0) \cap F^{-1}(y)$ and all $k \geq 1$. Consequently, using the definition of λ_k, we obtain

$$(1-\eta)^2 \sum_{k=0}^{\infty} \theta_k \frac{||F_0(x_k)||^4}{||F'(x_k)^* F_0(x_k)||^2} = (1-\eta) \sum_{k=0}^{\infty} \theta_k \lambda_k ||F_0(x_k)||^2 < \infty, \tag{12}$$

which, in particular, proves (10).

If $\sum \theta_k \lambda_k < \infty$ then $\sum ||x_k - x_{k+1}|| < \infty$ (due to (8a) and **A1**)) and (x_k) is a Cauchy sequence.

Suppose that $\sum \theta_k \lambda_k = \infty$. It follows from (12) that $\liminf ||F_0(x_k)|| = 0$. Since we are in Case II, the sequence $(||F_0(x_k)||)$ is strictly positive and there exists a subsequence (x_{ℓ_i}) satisfying

$$0 \leq k \leq \ell_i \Rightarrow ||F_0(x_k)|| \geq ||F_0(x_{\ell_i})||. \tag{13}$$

For all $k \in \mathbb{N}$ and $z \in B_\rho(x_0)$,

$$\begin{aligned}
||x_k - z||^2 &= ||x_{k+1} - z||^2 - ||x_k - x_{k+1}||^2 - 2\langle x_k - x_{k+1}, x_k - z \rangle \\
&\leq ||x_{k+1} - z||^2 - 2\langle x_k - x_{k+1}, \rangle \\
&= ||x_{k+1} - z||^2 + 2\theta_k \lambda_k \langle F'(x_k)^* F_0(x_k), x_k - z \rangle \\
&\leq ||x_{k+1} - z||^2 + 8\lambda_k (||F_0(x_k)||^2 + ||F_0(x_k)|| ||F_0(z)||),
\end{aligned} \tag{14}$$

where the second equality follows from (8a) and the last inequality follows from Proposition 2.1, item 2, and the assumption $\eta < 1$. Thus, taking $z = x_{\ell_i}$ in (14), we obtain

$$||x_k - x_{\ell_i}||^2 \leq ||x_{k+1} - x_{\ell_i}||^2 + 16\lambda_k ||F_0(x_k)^2||, \text{ for } 0 \leq k < \ell_i.$$

Define $s_m = \sum_{k \geq m} \theta_k \lambda_k ||F_0(x_k)||^2$. It follows from (12) that $\lim_{m \to \infty} s_m = 0$. If $0 \leq k < \ell_i$, by adding the above inequality for $j = k$, $k+1$, ..., $\ell_i - 1$, we get

$$||x_k - x_{\ell_i}||^2 \leq 16 \sum_{j=k}^{\ell_i - 1} \lambda_j ||F_0(x_j)||^2 \leq 16 s_k.$$

Now, take $k < j$. There exists $\ell_i > j$. Since $s_k > s_j$,

$$||x_k - x_j|| \leq ||x_k - x_{\ell_i}|| + ||x_j - x_{\ell_i}|| \leq 4\sqrt{s_k} + 4\sqrt{s_j} \leq 8\sqrt{s_k}.$$

Therefore, (x_k) is a Cauchy sequence and converges to some element $\bar{x} \in \overline{B_\rho(x_0)}$. □

Proof. (Theorem 3.4) It follows from the assumptions of statement (a), from Theorem 3.3, and from **A1** that (x_k) converges to some $\bar{x} \in B_\rho(x_0)$ and that

$$0 = \lim_{k \to \infty} \frac{||F_0(x_k)||^4}{||F'(x_k)^* F_0(x_k)||^2} \geq \limsup_{k \to \infty} \frac{||F_0(x_k)||^2}{C^2}.$$

Assertion (a) follows now from Proposition 2.2.

To prove item (b), first use Cauchy–Schwarz inequality to obtain

$$0 < \frac{||F'(x_k)^* F_0(x_k)||^4}{||F_0(x_k)||^2 ||F'(x_k) F'(x_k)^* F_0(x_k)||^2} \leq \frac{||F'(x_k)^* F_0(x_k)||^4}{\langle F_0(x_k), F'(x_k) F'(x_k)^* F_0(x_k) \rangle^2} = 1$$

Therefore, $0<\theta_k \leq (1-\eta)^{-1}<2$ for all k and it follows from Theorem 3.3, the definition of θ_k and from **A1** that (x_k) converges to some $\bar{x} \in B_\rho(x_0)$ and that

$$0 = \lim_{k\to\infty} \frac{||F_0(x_k)||^2 ||F'(x_k)^*F_0(x_k)||^2}{||F'(x_k)F'(x_k)^*F_0(x_k)||^2} \geq \limsup_{k\to\infty} \frac{||F_0(x_k)||^2}{C^2}.$$

Assertion (b) follows now from Proposition 2.2.

It follows from the assumptions of statement (c) that $0<\theta_k<(1-\eta)^{-1}<2$. From this point on, the proof of statement (c) is analogous to the proof of statement (b).

It follows from the assumptions of statement (d) that $0<\theta_k<2$. As before, the proof of statement (d) is analogous to the proof of statement (b).　□

Remark 3.5. *The argument used to establish strong convergence of sequence (x_k) in the proof of Theorem 3.3 is inspired by the technique used in [3, Theorem 2.3] to prove an analog result for the nonlinear Landweber iteration. Both proofs rely on a Cauchy sequence argument (it is necessary to prove that (x_k) is a Cauchy sequence). In [3], given $j \geq k$ arbitrarily large, an element $j \geq l \geq k$ is chosen with a minimal property (namely, $||F_0(x_l)|| \leq ||F_0(x_i)||$, for $k \leq i \leq j$). In the proof of Theorem 3.3, the auxiliary indexes ℓ_i defined in (13) play a similar role. These indexes are also chosen according to a minimizing property, namely, the subsequence $(||F_0(x_{\ell_j})||)$ is monotone non-increasing.*

4. Convergence analysis: Noisy data

In this section, we analyze the family of relaxed projected Landweber methods in the noisy data case and investigate convergence properties. We assume that only noisy data $y^\delta \in Y$ satisfying (1) are available, where the noise level $\delta>0$ is known. Recall that to simplify the presentation we are using notation (3), i.e. $F_\delta(x) = F(x)-y^\delta$.

Since $F_0(\cdot) = F(\cdot)-y$ is not available, one cannot compute the projection onto H_x (defined in Section 3). Define, instead, for each $x \in B_\rho(x_0)$, the set

$$H_x^\delta := \{z \in X | \langle z-x, F'(x)^*F_\delta(x)\rangle \leq \\ \leq -||F_\delta(x)||((1-\eta)||F_\delta(x)||-(1+\eta)\delta)\}. \tag{15}$$

Next we prove a "noisy" version of the separation Lemma 3.1: H_x^δ contains all exact solutions of $F(x)=y$ (within $B_\rho(x_0)$) and, if the residual $||F_\delta(x)||$ is above the threshold $(1+\eta)(1-\eta)^{-1}\delta$, then H_x^δ does not contain x.

Lemma 4.1 (Separation): *Suppose that **A1** and **A2** hold. If $x \in B_\rho(x_0)$, then*

$$0 \geq ||F_\delta(x)|| [(1-\eta)||F_\delta(x)|| - (1+\eta)\delta] + \langle x^\star - x, F'(x)^* F_\delta(x) \rangle, \qquad (16)$$

for all $x^\star \in B_\rho(x_0) \cap F^{-1}(y)$. Consequently, $B_\rho(x_0) \cap F^{-1}(y) \subset H_x^\delta$.

Proof. Indeed, for $x^\star \in B_\rho(x_0) \cap F^{-1}(y)$ we have

$$\begin{aligned}
\langle F'(x)^* F_\delta(x), x^\star - x \rangle &= \langle F_\delta(x), F'(x)(x^\star - x) \rangle \\
&= \langle F_\delta(x), F_\delta(x) + F'(x)(x^\star - x) \rangle - ||F_\delta(x)||^2 \\
&= \langle F_\delta(x), F_0(x) + F'(x)(x^\star - x) \rangle + \langle F_\delta(x), y - y^\delta \rangle ||F_\delta(x)||^2 \\
&\leq ||F_\delta(x)|| \eta ||F_0(x)|| + ||F_\delta(x)|| \delta - ||F_\delta(x)||^2
\end{aligned}$$

where, the first inequality follows from Cauchy–Schwarz inequality and (5). Since $||F_0(x)|| \leq ||F_\delta(x)|| + \delta$,

$$\langle F'(x)^* F_\delta(x), x^\star - x \rangle \leq \eta ||F_\delta(x)|| (||F_\delta(x)|| + \delta) + ||F_\delta(x)|| \delta - ||F_\delta(x)||^2$$

which is equivalent to (16). □

Since $||F_\delta(x)|| > (1+\eta)(1-\eta)^{-1}\delta$ is sufficient for separation of x from $F^{-1}(y)$ in $B_\rho(x_0)$ via H_x^δ, this condition also guarantees $F'(x)^* F_\delta(x) \neq 0$.

The iteration formula for the family of relaxed projection Landweber methods in the noisy data case is given by

$$x_{k+1}^\delta := x_k^\delta - \theta_k \frac{p_\delta(||F_\delta(x_k^\delta)||)}{||F'(x_k^\delta)^* F_\delta(x_k^\delta)||^2} F'(x_k^\delta)^* F_\delta(x_k^\delta), \theta_k \in (0, 2), \qquad (17)$$

where

$$p_\delta(t) := t((1-\eta)t - (1+\eta)\delta) \qquad (18)$$

and the initial guess $x_0^\delta \in X$ is chosen according to **A1**. Again, the *PLW method* (for inexact data) is obtained by taking $\theta_k = 1$, which amounts to define x_{k+1}^δ as the orthogonal projection of x_k^δ onto $H_{x_k^\delta}^\delta$. On the other hand, the relaxed variants, which use $\theta_k \in (0, 2)$, correspond to setting x_{k+1}^δ as a relaxed projection of x_k^δ onto $H_{x_k^\delta}^\delta$.

Let

$$\tau > \frac{1 + \eta}{1 - \eta}. \qquad (19)$$

The computation of the sequence (x_k^δ) should be stopped at the index $k_*^\delta \in \mathbb{N}$ defined by the discrepancy principle

$$k_*^\delta := \max\{k \in \mathbb{N}; ||F_\delta(x_j^\delta)|| > \tau\delta, j = 0, 1, ..., k-1\}. \qquad (20)$$

Notice that if $||F_\delta(x_k^\delta)|| > \tau\delta$, then $||F'(x_k^\delta)^* F_\delta(x_k^\delta)|| \neq 0$. This fact is a consequence of Proposition 2.1, item 3, since F_δ also satisfies **A1** and **A2**. Consequently, iteration (17) is well defined for $k = 0, ..., k_*^\delta$.

The next two results have interesting consequences. From Proposition 4.2, we conclude that (x_k^δ) does not leave the ball $B_\rho(x_0)$ for $k = 0, ..., k_*^\delta$. On the other hand, it follows from Theorem 4.3 that the stopping index k_*^δ is finite, whenever $\delta > 0$.

Proposition 4.2. *Let **A1**–**A3** hold true and θ_k be chosen as in (17). If $x_k^\delta \in B_\rho(x_0)$ and $||F_\delta(x_k^\delta)|| > \tau\delta$, then*

$$||x^\star - x_k^\delta||^2 \geq ||x^\star - x_{k+1}^\delta||^2 + \theta_k(2 - \theta_k)\left(\frac{p_\delta(||F_\delta(x_k^\delta)||)}{||F\prime(x_k^\delta)^* F_\delta(x_k^\delta)||}\right)^2,$$

for all $x^\star \in B_\rho(x_0) \cap F^{-1}(y)$.

Proof. If $x_k^\delta \in B_\rho(x_0)$ and $||F_\delta(x_k^\delta)|| > \tau\delta$, then x_{k+1}^δ is a relaxed orthogonal projection of x_k^δ onto $H_{x_k^\delta}^\delta$ with a relaxation factor θ_k. The conclusion follows from this fact, Lemma 4.1, the iteration formula (17), and elementary properties of over/under relaxed orthogonal projections.　□

Theorem 4.3. *If **A1**–**A3** hold true, then the sequences (x_k^δ), (θ_k) as specified in (17) (together with the stopping criterion (20)) are well defined and*

$$x_k \in B_{\rho/2}(x^\star) \subset B_\rho(x_0), \forall k \leq k_*^\delta.$$

Moreover, if $\theta_k \in [a, b] \subset (0, 2)$ for all $k \leq k_*^\delta$, then this stopping index k_*^δ defined in (20) is finite.

Proof. The proof of the first statement is similar to the one in Theorem 3.3.

To prove the second statement, first observe that since $\theta_k \in [a, b]$, $\theta_k(2 - \theta_k) \geq a(2 - b) > 0$. Thus, it follows from Proposition 4.2 that for any $k < k_*^\delta$

$$||x^\star - x_0^\delta||^2 \geq a(2 - b)\sum_{j=0}^{k}\left(\frac{p_\delta(||F_\delta(x_k^\delta)||)}{||F\prime(x_k^\delta)^* F_\delta(x_k^\delta)||}\right)^2$$

$$\geq \frac{a(2 - b)}{C^2}\sum_{j=0}^{k}\left(\frac{p_\delta(||F_\delta(x_k^\delta)||)}{||F_\delta(x_k^\delta)||}\right)^2.$$

Observe that, if $t > \tau\delta$, then

$$\frac{p_\delta(t)}{t} = (1 - \eta)t - (1 + \eta)\delta > \left[\tau - \frac{1 + \eta}{1 - \eta}\right](1 - \eta)\delta =: h > 0.$$

Therefore, for any $k < k_*^\delta$
$$||x^\star - x_0^\delta||^2 \geq \frac{a(2 - b)}{C^2}(k + 1)h^2,$$

so that k_*^δ is finite.　□

It is worth noticing that the Landweber method for noisy data [2, Chap. 11] (which requires $\eta < 1/2, C \leq 1$ in **A1–A2**) using the discrepancy principle (20) with

$$\tau > 2 \frac{1 + \eta}{1 - 2\eta} > \frac{1 + \eta}{1 - \eta},$$

corresponds to the PLW method, analyzed in Theorem 4.3, with

$$0 < \frac{p_\delta(\tau\delta)}{\rho^2} \leq \theta_k = \frac{||F'(x_k^\delta)^* F_\delta(x_k^\delta)||^2}{p_\delta(||F_\delta(x_k^\delta)||)} \leq \frac{\tau}{(1-\eta)\tau - (1+\eta)} < 2$$

(Here the second inequality follows from **A3** and the third inequality follows from Lemma 4.1). Consequently, in the noisy data case, the convergence analysis for the PLW method encompasses the Landweber iteration (under the TCC condition) as a particular case.

In the next theorem we discuss a stability result, which is an essential tool to prove the last result of this section, namely Theorem 4.5 (semi-convergence of the PLW method). Notice that this is the first time where the strong Assumption **A4** is needed in the text.

Theorem 4.4. *Let **A1–A4** hold true. For each fixed $k \in \mathbb{N}$, the element x_k^δ, computed after kth-iterations of any method within the family of methods in (17), depends continuously on the data y^δ.*

Proof. From (19), **A1**, **A4** and Theorem 4.3, it follows that the mapping $\varphi : D(\varphi) \to X$ with

$$D(\varphi) := \{(x, y^\delta, \delta) | x \in D(F); \delta > 0; ||y^\delta - y|| \leq \delta; F'(x)^*(F(x) - y^\delta) \neq 0\},$$
$$\varphi(x, y^\delta, \delta) := x - \frac{p_\delta(||F(x) - y^\delta||)}{||F'(x)^*(F(x) - y^\delta)||^2} F'(x)^*(F(x) - y^\delta)$$

is continuous on its domain of definition. Therefore, whenever the iterate $x_k^\delta = (\varphi(\cdot, y^\delta, \delta))^k (x_0)$ is well defined, it depends continuously on (y^δ, δ). □

Theorem 4.4 together with Theorems 3.3 and 3.4 are the key ingredients in the proof of Theorem 4.5, which guarantees that the stopping rule (20) renders the PLW iteration a regularization method. The proof of Theorem 4.5 uses classical techniques from the analysis of Landweber-type iterative regularization techniques (see, e.g., [2, Theorem 11.5] or [4, Theorem 2.6]) and thus is omitted.

Theorem 4.5. *Let **A1** – **A4** hold true, $\delta_j \to 0$ as $j \to \infty$, and $y_j := y^{\delta_j} \in Y$ be given with $||y_j - y|| \leq \delta_j$. If the PLW iteration (17) is stopped with $k_*^j := k_*^{\delta_j}$ according to the discrepancy (20), then $(x_{k_*^j}^{\delta_j})$ converges strongly to a solution $\bar{x} \in B_\rho(x_0)$ of $F(x) = y$ as $j \to \infty$.*

It is immediate to verify that the results in Theorem 4.5 extend to any method within the family of relaxed projection Landweber methods (17).

5. Numerical experiments

In this section, we present numerical experiments for the iterative methods derived in previous sections. The PLW method is implemented for solving an exponentially ill-posed inverse problem related to the Dirichlet to Neumann operator and its performance is compared against the benchmark methods LW and SD.

5.1. Description of the mathematical model

We briefly introduce a model which plays a key role in inverse doping problems with current flow measurements, namely the two-dimensional *linearized stationary bipolar model close to equilibrium*.

This mathematical model is derived from the drift diffusion equations by linearizing the Voltage–Current (VC) map at $U \equiv 0$ [18, 19], where the function $U = U(x)$ denotes the applied potential to the semiconductor device.[1] Additionally, we assume that the electron mobility $\mu_n(x) = \mu_n > 0$ as well as the hole mobility $\mu_p(x) = \mu_p > 0$ are constant and that no recombination-generation rate is present [20, 21]. Under the above assumptions, the Gateaux derivative of the VC-map Σ_C at the point $U = 0$ in the direction $h \in H^{3/2}(\partial\Omega_D)$ is given by

$$\Sigma_C'(0)h = \mu_n e^{V_{bi}}\hat{u}_\nu - \mu_p e^{-V_{bi}}\hat{v}_\nu \in H^{1/2}(\Gamma_1), \tag{21}$$

where the concentrations of electrons and holes $(\hat{u}, \hat{v})$ solve[2]

$$\text{div}(\mu_n e^{V^0}\nabla\hat{u}) = 0 \ \text{ in } \ \Omega \tag{22a}$$
$$\text{div}(\mu_p e^{-V^0}\nabla\hat{v}) = 0 \ \text{ in } \ \Omega \tag{22b}$$
$$\hat{u} = -\hat{v} = -h \ \text{ on } \ \partial\Omega_D \tag{22c}$$
$$\nabla\hat{u} \cdot \nu = \nabla\hat{v} \cdot \nu = 0 \ \text{ on } \ \partial\Omega_N \tag{22d}$$

and the potential V^0 is the solution of the thermal equilibrium problem

$$\lambda^2 \Delta V^0 = e^{V^0} - e^{-V^0} - C(x) \ \text{ in } \ \Omega \tag{23a}$$
$$V^0 = V_{bi}(x) \ \text{ on } \ \partial\Omega_D \tag{23b}$$
$$\nabla V^0 \cdot \nu = 0 \ \text{ on } \ \partial\Omega_N. \tag{23c}$$

Here, $\Omega \subset \mathbb{R}^2$ is a domain representing the semiconductor device; the boundary of Ω is divided into two nonempty disjoint parts: $\partial\Omega = \overline{\partial\Omega_N} \cup \overline{\partial\Omega_D}$. The Dirichlet boundary part $\partial\Omega_D$ models the Ohmic

---

[1]This simplification is motivated by the fact that, due to hysteresis effects for large applied voltage, the VC-map can only be defined as a single-valued function in a neighborhood of $U = 0$.
[2]These concentrations are written in terms of the Slotboom variables [15].

contacts, where the potential V as well as the concentrations $\hat{u}$ and $\hat{v}$ are prescribed; the Neumann boundary part $\partial\Omega_N$ corresponds to insulating surfaces, thus zero current flow and zero electric field in the normal direction are prescribed; the Dirichlet boundary part splits into $\partial\Omega_D = \Gamma_0 \cup \Gamma_1$, where the disjoint curves Γ_i, $i = 0$, 1, correspond to distinct device contacts (differences in $U(x)$ between segments Γ_0 and Γ_1 correspond to the applied bias between these two contacts). Moreover, V_{bi} is a given logarithmic function [18].

The piecewise constant function $C(x)$ is the *doping profile* and models a preconcentration of ions in the crystal, so $C(x) = C_+(x) - C_-(x)$ holds, where C_+ and C_- are (constant) concentrations of negative and positive ions, respectively.

In those subregions of Ω in which the preconcentration of negative ions predominate (P-regions), we have $C(x) < 0$. Analogously, we define the N-regions, where $C(x) > 0$ holds. The boundaries between the P-regions and N-regions (where C changes sign) are called *pn-junctions*; it's determination is a strategic non-destructive test [20, 21].

5.2. The inverse doping problem

The inverse problem we are concerned with consists of determining the doping profile function C in (23) from measurements of the linearized VC-map $\Sigma'_C(0)$ in (21), under the assumption $\mu_p = 0$ (the so-called *linearized stationary unipolar model close to equilibrium*). Notice that we can split the inverse problem into two parts:

1. Define the function $a(x) := e^{V^0(x)}, x \in \Omega$, and solve the parameter identification problem

$$\text{div}(\mu_n a(x)\nabla\hat{u}) = 0 \ \text{ in } \ \Omega \quad \hat{u} = -U(x) \ \text{ on } \ \partial\Omega_D \quad \nabla\hat{u}\cdot\nu = 0 \ \text{ on } \ \partial\Omega_N.$$
(24)

for $a(x)$, from measurements of

$$\left[\Sigma'_C(0)\right](U) = (\mu_n a(x)\hat{u}_\nu)|_{\Gamma_1}.$$

2. Evaluate the doping profile
$$C(x) = a(x) - a^{-1}(x) - \lambda^2\Delta(\ln a(x)), \ x \in \Omega.$$

Since the evaluation of C from $a(x)$ can be explicitly performed in a stable way, we shall focus on the problem of identifying the function parameter $a(x)$ in (24). Summarizing, the inverse doping profile problem in the linearized stationary unipolar model (close to equilibrium) reduces to the identification of the parameter function $a(x)$ in (24) from measurements of the Dirichlet-to-Neumann map $\Lambda_a : H^{1/2}(\partial\Omega_D) \ni U \mapsto (\mu_n a(x)\hat{u}_\nu)|_{\Gamma_1} \in H^{-1/2}(\Gamma_1)$.
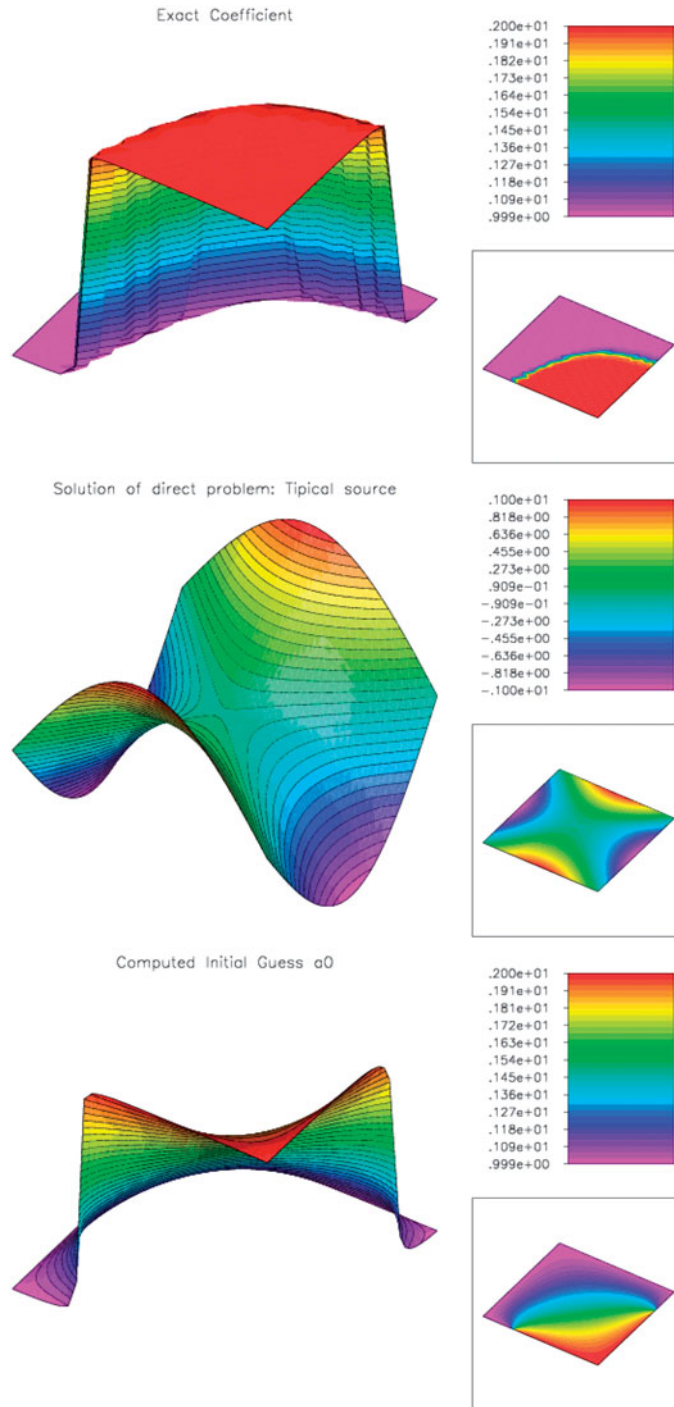
Figure 1. First experiment: setup of the problem. **Top:** The parameter $a^\star(x)$ to be identified; **Center:** Voltage source $U(x)$ (Dirichlet boundary condition at $\partial\Omega$ for the DtN map) and the corresponding solution $\hat{u}$ of (24); **Bottom:** Initial guess $a_0(x)$ for the iterative methods PLW, LW, and SD.

In the formulation of the inverse problem we shall take into account some constraints imposed by the practical experiments, namely: (i) The voltage profile $U \in H^{1/2}(\partial\Omega_D)$ must satisfy $U|_{\Gamma_1} = 0$ (in practice, U is chosen to be piecewise constant on Γ_1 and to vanish on Γ_0); (ii) The identification of $a(x)$ has to be performed from a finite number of measurements, i.e. from the data $\{(U_i, \Lambda_a(U_i))\}_{i=1}^{N} \in [H^{1/2}(\Gamma_0) * H^{-1/2}(\Gamma_1)]^N$.

In what follows we take $N=1$, i.e. identification of $a(x)$ from a single experiment. Thus, we can write this particular inverse doping problem within the abstract framework of (2)

$$F(a) = \Lambda_a(U) =: y, \tag{25}$$

where U is a fixed voltage profile satisfying the above assumptions, $X := L^2(\Omega) \supset D(F) := \{a \in L^\infty(\Omega); \ 0 < a_m \le a(x) \le a_M, \text{ a.e. in } \Omega\}$ and $Y := H^{1/2}(\Gamma_1)$. The operator F above is known to be continuous [18].

5.3. First experiment: the Calderon setup

In this subsection, we consider the special setup $\Gamma_1 = \partial\Omega_D = \partial\Omega$ (i.e., $\Gamma_0 = \partial\Omega_N = \varnothing$). Up to now, it is not known whether the map F satisfies the TCC. However,

1. the map $a \mapsto u$ (solution of (24)) satisfies the TCC with respect to the $H^1(\Omega)$ norm [4];
2. it was proven in [22] that the discretization of the operator F in (25) using the finite element method (and basis functions constructed by a Delaunay triangulation) satisfies the TCC (5).

Therefore, the analytical convergence results of the previous sections do apply to finite-element discretizations of (25) in this special setup. Moreover, item 1 suggests that $H^1(\Omega)$ is a good choice of parameter space for TCC-based reconstruction methods. Motivated by this fact, the setup of the numerical experiments presented in this subsection is chosen as follows:

- The domain $\Omega \subset \mathbb{R}^2$ is the unit square $(0,1) * (0,1)$ and the above mentioned boundary parts are $\Gamma_1 = \partial\Omega_D := \partial\Omega$, $\Gamma_0 = \partial\Omega_N := \varnothing$.
- The parameter space is $H^1(\Omega)$ and the function $a^\star(x)(x)$ to be identified is shown in Figure 1.
- The fixed Dirichlet input for the DtN map (24) is the continuous function $U : \partial\Omega \to \mathbb{R}$ defined by

$$U(x,0) = U(x,1) := \sin(\pi x), \quad U(0,y) = U(1,y) := -\sin(\pi y)$$

(in Figure 1, $U(x)$ and the corresponding solution $\hat{u}$ of (24) are plotted).

- The TCC constant η in (5) is not known for this particular setup. In our computations, we used the value $\eta = 0.45$ which is in agreement with **A2**. (Note that the convergence analysis of the PLW method requires $\eta < 1$ while the nonlinear LW method requires the TCC with $\eta < 0.5$ [4, Assumption (2.4)]. The above choice allows the comparison of both methods.)
- The "exact data" y in (25) is obtained by solving the direct problem (24) using a finite element type method and adaptive mesh refinement (approximately 131.000 elements). In order to avoid inverse crimes, a coarser grid (with approximately 33.000 elements) was used in the finite element method implementation of the iterative methods.
- In the numerical experiment with noisy data, artificially generated (random) noise of 2% was added to the exact data y in order to generate the noisy data y^δ. For the verification of the stopping rule (20) we assumed exact knowledge of the noise level and chose $\tau = 3$ in (19), which is in agreement with the above choice for η.

Remark 5.1. (Choosing the initial guess): The initial guess $a_0(x)$ used for all iterative methods is presented in Figure 1. According to **A1–A3**, $a_0(x)$ has to be sufficiently close to $a^\star(x)$ (otherwise the PLW method may not converge). With this in mind, we choose $a_0(x)$ as the solution of the Dirichlet boundary value problem

$$\Delta a_0 = 0, \text{ in } \Omega, \qquad a_0(x) = U(x), \text{ at } \partial\Omega.$$

This choice is an educated guess that incorporate the available *a priori* knowledge about the exact solution $a^\star(x)$, namely: $a_0 \in H^1(\Omega)$ and $a_0(x) = a^\star(x)$ at $\partial\Omega_D$. Moreover, $a_0 = \operatorname{argmin}\{||\nabla a||^2_{L^2(\Omega)} \mid a \in H^1(\Omega), a^\Omega_\partial(x) = a^\star_{\partial\Omega}(x)\}$.

Remark 5.2 (Computing the iterative step): The computation of the kth-step of the PLW method (see (8)) requires the evaluation of $F'(a_k)^* F_0(a_k)$. According to [18], for all test functions $v \in H^1_0(\Omega)$ it holds

$$\langle F'(a_k)^* F_0(a_k), v\rangle_{L^2(\Omega)} = \langle F_0(a_k), F'(a_k)v\rangle_{L^2(\partial\Omega)} = \langle F_0(a_k), V\rangle_{L^2(\partial\Omega)},$$

where $F'(a_k)^*$ stands for the adjoint of $F'(a_k)$ in $L^2(\Omega)$, and $V \in H^1(\Omega)$ solves

$$-\nabla \cdot (a_k(x)\nabla V) = \nabla \cdot (v\nabla F(a_k)), \text{ in } \Omega, \qquad V = 0, \text{ at } \partial\Omega.$$

Furthermore, in [18] it is shown that for all $\psi \in L^2(\partial\Omega)$ and $v \in H^1_0(\Omega)$

$$\langle F'(a_k)^*\psi, v\rangle_{L^2(\Omega)} = \langle \psi, V\rangle_{L^2(\partial\Omega)} = \langle \nabla\Psi \cdot \nabla u_k, v\rangle_{L^2(\Omega)}, \tag{26}$$

where Ψ, $u_k \in H^1(\Omega)$ solve

$$-\nabla \cdot (a_k(x)\nabla\Psi) = 0, \text{ in } \Omega, \qquad \Psi = \psi, \text{ at } \partial\Omega \tag{27a}$$
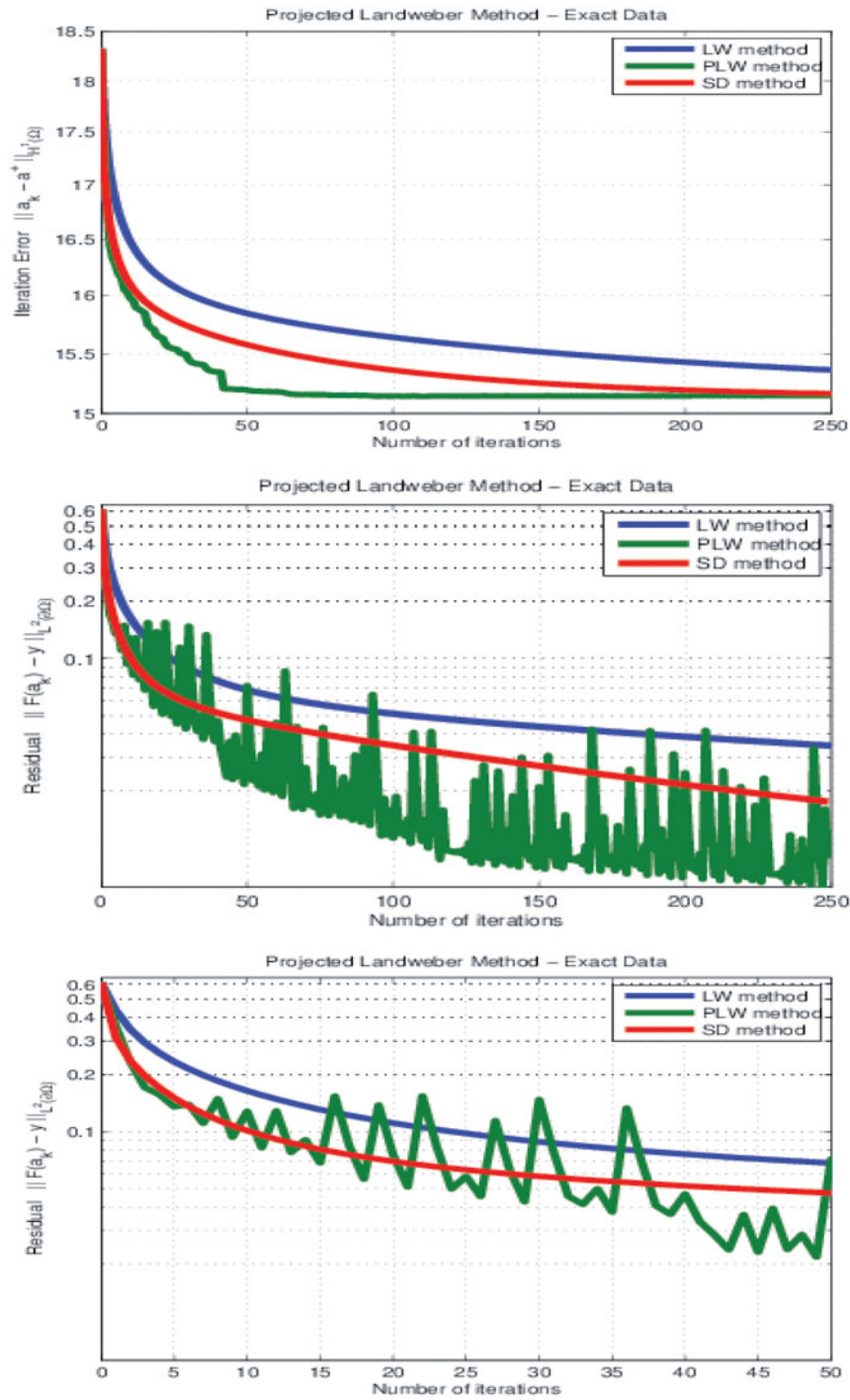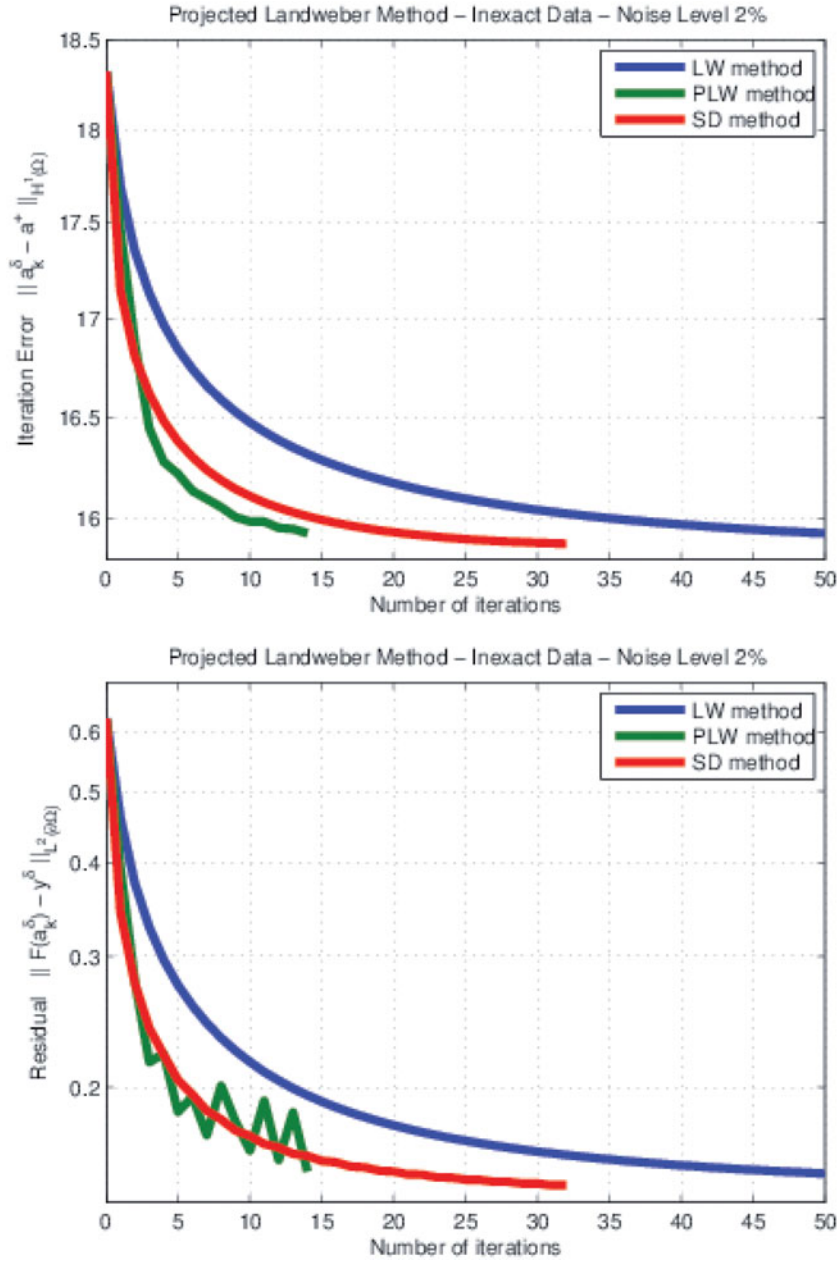
Figure 2. First experiment: example with exact data. The PLW method (GREEN) is compared with the LW method (BLUE) and with the SD method (RED); **Top:** Iteration error $||a_k - a^\star||_{H^1(\Omega)}$; **Middle:** Residual $||F(a_k) - y||_{L^2(\partial\Omega)}$; **Bottom:** Residual, detail of the first 50 iterations.

Figure 3. First experiment: example with noisy data. The PLW method (GREEN) is compared with the LW method (BLUE) and with the SD method (RED); **Top:** Iteration error $||a_k^\delta - a^\star||_{H^1(\Omega)}$; **Bottom:** Residual $||F(a_k^\delta) - y^\delta||_{L^2(\partial\Omega)}$.

$$-\nabla \cdot (a_k(x)\nabla u_k) = 0, \text{ in } \Omega, \qquad u_k = U(x), \text{ at } \partial\Omega. \qquad (27b)$$

respectively. A direct consequence of (26), (27) is the variational identity

$$\langle F'(a_k)^* F_0(a_k), v \rangle_{L^2(\Omega)} = \langle \nabla\Psi \cdot \nabla u_k, v \rangle_{L^2(\partial\Omega)}, \forall v \in H_0^1(\Omega),$$

where Ψ solves (27a) with $\psi = F_0(a_k)$.

Notice that $\nabla\Psi \cdot \nabla u_k$ is the adjoint, in $L^2(\Omega)$, of $F'(a_k)$ applied to $F_0(a_k)$. We need to apply to $F_0(a_k)$, instead, the adjoint of $F'(a_k)$ in $H^1(\Omega)$. That is, we need to compute

$$F'(a_k)^* F_0(a_k) = W_k \in H_0^1(\Omega),$$

Where, W_k is the Riesz vector satisfying $\langle W_k, v \rangle_{H^1(\Omega)} = \langle \nabla\Psi \cdot \nabla u_k, v \rangle_{L^2(\Omega)}$, for all $v \in H^1(\Omega)$. A direct calculation yields

$$(I-\Delta)W_k = \nabla\Psi \cdot \nabla u_k, \text{ in } \Omega, \qquad W_k = 0, \text{ at } \partial\Omega.$$

Within this setting, the PLW iteration (8) becomes

$$a_{k+1} := a_k - (1-\eta)\frac{||F_0(a_k)||_{L^2(\Omega)}^2}{||W_k||_{H^1(\Omega)}^2} W_k.$$

The iterative steps of the benchmark iterations LW and SD (implemented here for the sake of comparison) are computed also using the adjoint of $F'(\cdot)$ in H^1. Notice that a similar argumentation can be derived in the noisy data case (see (17)).

For solving the elliptic PDE's described above, needed for the implementation of the iterative methods, we used the package PLTMG [23] compiled with GFORTRAN-4.8 in a INTEL(R) Xeon(R) CPU E5-1650 v3 (Santa Clara, CA).

First example: Problem with exact data.

Evolution of both iteration error and residual is shown in Figure 2. The PLW method (GREEN) is compared with the LW method (BLUE) and with the SD method (RED). For comparison purposes, if one decides to stop iterating when $||F_0(a_k)|| < 0.025$ is satisfied, the PLW method needs only 43 iterations, while the SD method requires 167 iterative steps and the LW method required more than 500 steps.

Second example: Problem with noisy data.

Evolution of both iteration error and residual is shown in Figure 3. The PLW method (GREEN) is compared with the LW method (BLUE) and with the SD method (RED). The stop criteria (20) is reached after 14 steps of the PLW iteration, 32 steps for the SD iteration, and 56 steps for the LW iteration.

5.4. Second experiment: The semiconductor setup

In this paragraph, we consider the more realistic setup (in agreement with the semiconductor models in Subsection 5.1) with $\partial\Omega_D \not\subseteq \partial\Omega$, and $\Gamma_0 \neq \varnothing, \partial\Omega_N \neq \varnothing$.

In this experiment, we have: (i) The voltage profile $U \in H^{1/2}(\partial\Omega_D)$ satisfies $U|_{\Gamma_1} = 0$; (ii) As in the previous experiment, the identification of $a(x)$ is performed from a single measurement. To the best of our knowledge,
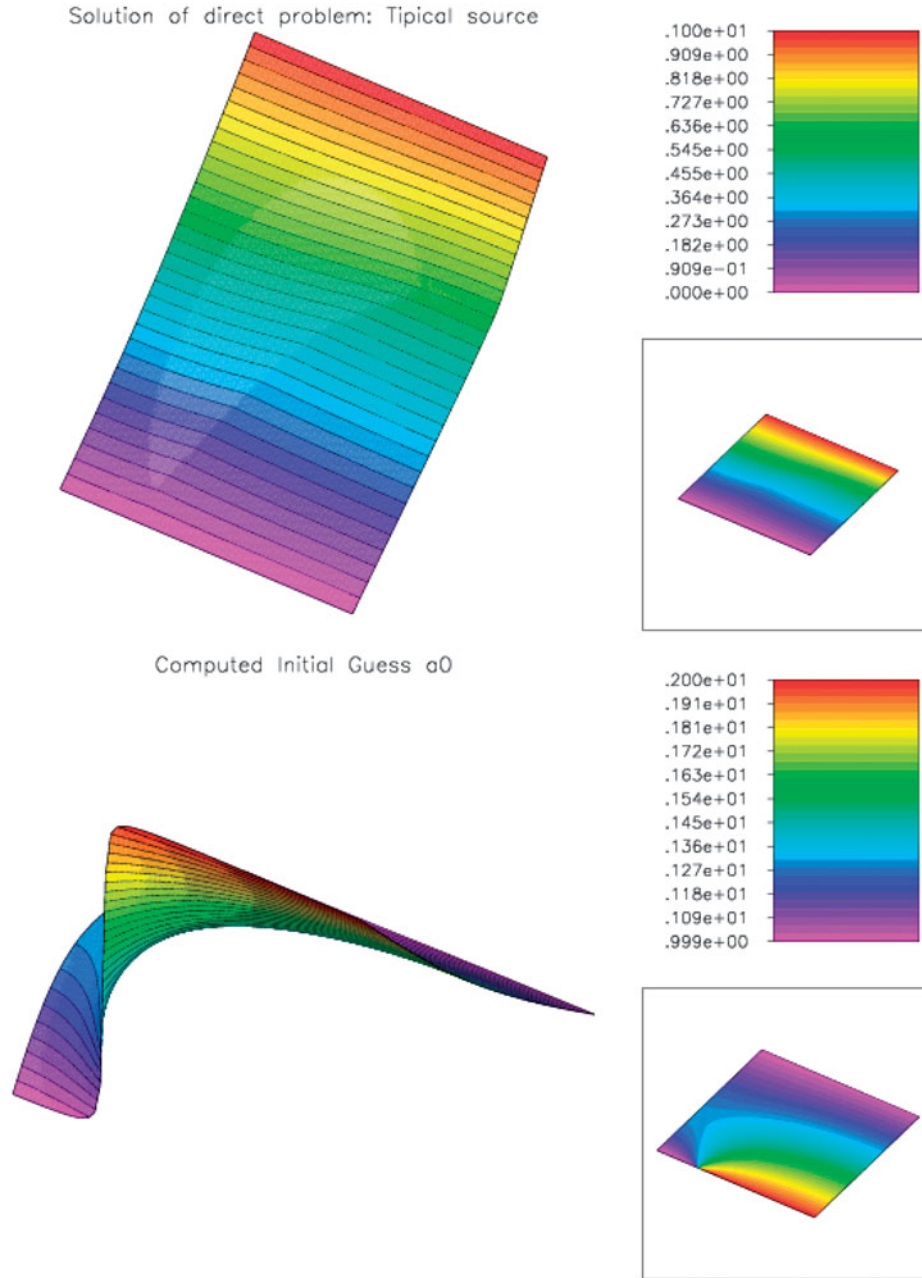
Solution of direct problem: Tipical source

Computed Initial Guess a0

Figure 4. Second experiment: setup of the problem. **Top:** Voltage source $U(x)$ (Dirichlet boundary condition at $\partial\Omega_D$ for the DtN map) and the corresponding solution $\hat{u}$ of (24); **Bottom:** Initial guess $a_0 \in H^1(\Omega)$ satifying $a_0(x) = U(x)$ at $\partial\Omega_D$ and $\nabla a_0(x) \cdot \nu(x) = 0$ at $\partial\Omega_N$.

within this setting, Assumptions **A1**–**A3** were not yet established for the operator F in (25) and its discretizations. Therefore, although the operator F is continuous [23], it is still unclear whether the analytical convergence results of the previous sections hold here.
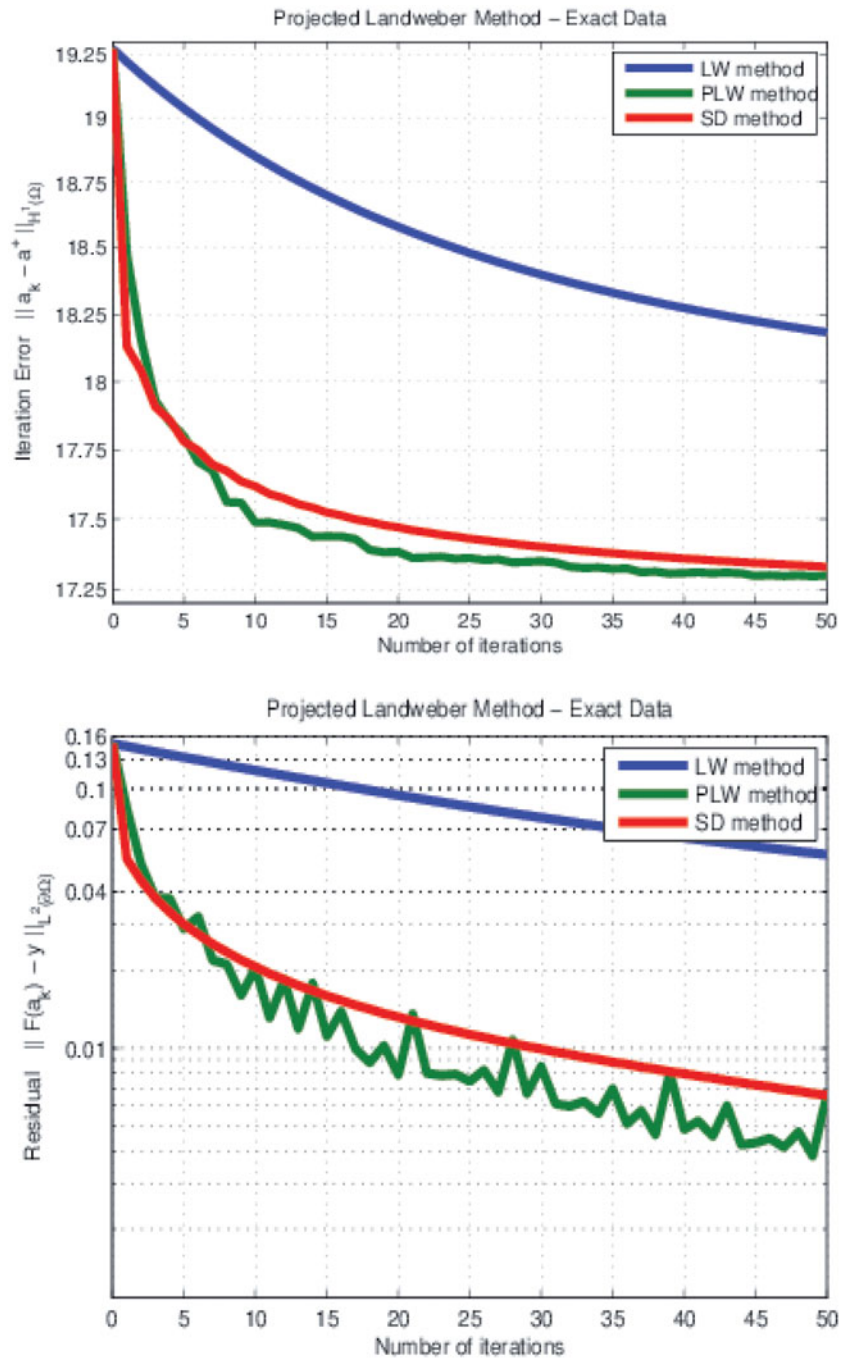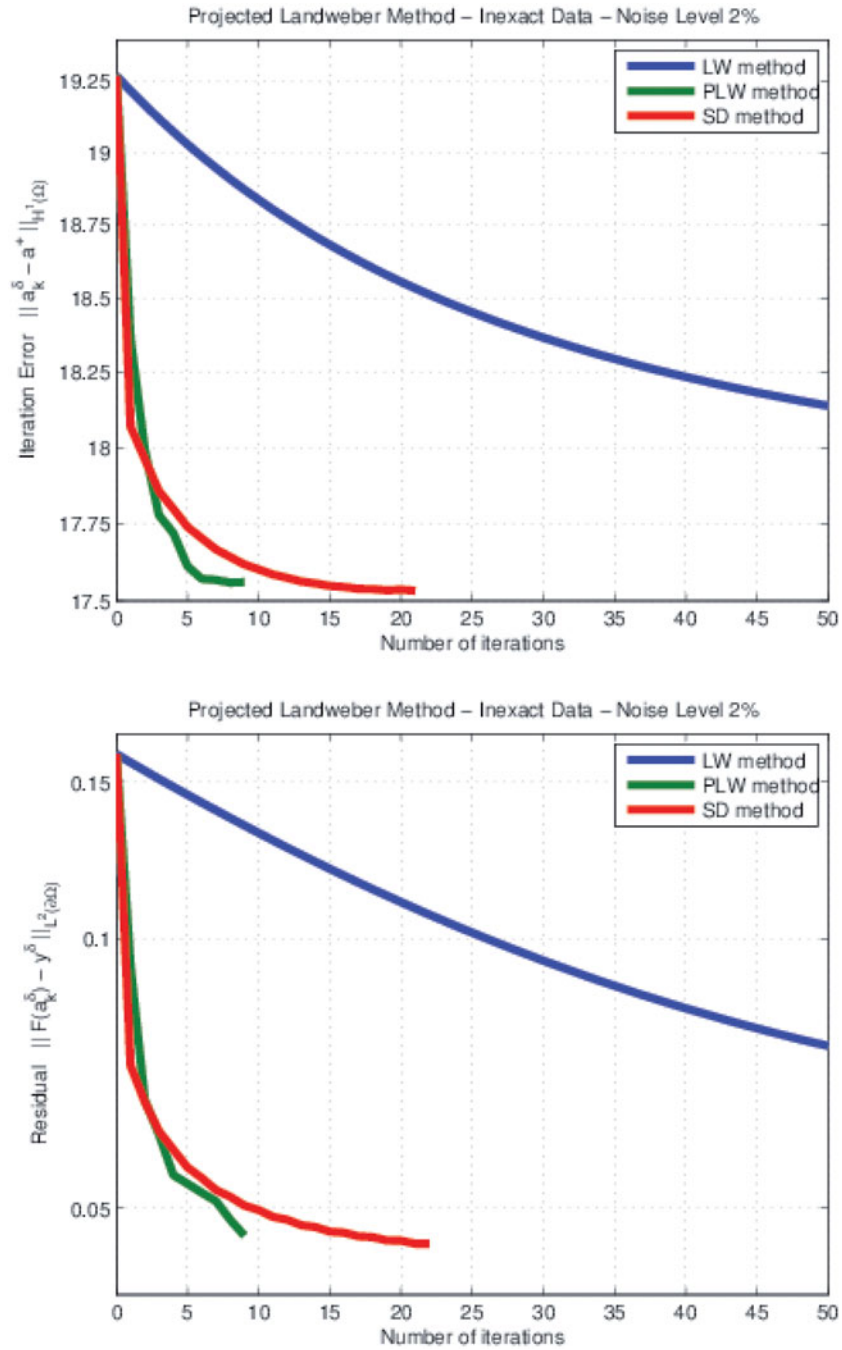
Figure 5. Second experiment: example with exact data. The PLW method (GREEN) is compared with the LW method (BLUE) and with the SD method (RED); **Top:** Iteration error $||a_k - a^\star||_{H^1(\Omega)}$; **Bottom:** Residual $||F(a_k) - y||_{L^2(\Gamma_1)}$.

Figure 6. Second experiment: example with noisy data. The PLW method (GREEN) is compared with the LW method (BLUE) and with the SD method (RED); **Top:** Iteration error $||a_k^\delta - a^\star||_{H^1(\Omega)}$; **Bottom:** Residual $||F(a_k^\delta) - y^\delta||_{L^2(\partial\Omega)}$.

The setup of the numerical experiments presented in this section is the following:

- The elements listed below are the same as in the previous experiment:
- The domain $\Omega \subset \mathbb{R}^2$;
- The parameter space $H^1(\Omega)$ and the function $a^\star(x)$ to be identified;
- The computation of the "exact data" y in (25);
- The choice for the TCC constant η in (5) and for τ in (19);
- The level δ of artificially introduced noise;
- The procedure to generate the noisy data y^δ;
- The boundary parts mentioned in Subsection 5.1 are defined by $\partial \Omega_D := \Gamma_0 \cup \Gamma_1$, $\Gamma_1 := \{(x, 1); x \in (0, 1)\}$, $\Gamma_0 := \{(x, 0); x \in (0, 1)\}$, $\partial \Omega_N := \{(0, y); y \in (0, 1)\} \cup \{(1, y); y \in (0, 1)\}$ (in Figure 4(a and b), the boundary part Γ_1 corresponds to the lower left edge, while Γ_0 is the top right edge; the origin is on the upper right corner).
- The fixed Dirichlet input for the DtN map (24) is the piecewise constant function $U : \partial \Omega_D \to \mathbb{R}$ is defined by $U(x, 0) := 1$, and $U(x, 1) = 0$. In Figure 4(a), $U(x)$ and the corresponding solution $\hat{u}$ of (24) are plotted.
- The initial condition $a_0(x)$ used for all iterative methods is shown in Figure 4(b) and is given by the solution of the mixed boundary value problem

$$\Delta a_0(x) = 0, \text{ in } \Omega, \qquad a_0(x) = U(x), \text{ at } \partial \Omega_D, \qquad \nabla a_0 \cdot \nu = 0, \text{ at } \partial \Omega_N,$$

analogously as in Remark 5.1.
- The computation of the iterative-step of the PLW method is performed analogously as in Remark 5.2, namely

$$a_{k+1} := a_k - (1-\eta) \frac{||F_0(a_k)||^2_{L^2(\Omega)}}{||W_k||^2_{H^1(\Omega)}} W_k.$$

where the Riesz vector $W_k \in H^1(\Omega)$ solves

$$(I - \Delta) W_k = \nabla \Psi \cdot \nabla u_k, \text{ in } \Omega, \qquad W_k = 0, \text{ at } \partial \Omega_D, \qquad \nabla W_k \cdot \nu = 0, \text{ at } \partial \Omega_N,$$

and Ψ, u_k solve

$$-\nabla \cdot (a_k(x) \nabla \Psi) = 0, \text{ in } \Omega, \qquad \Psi = F_0(a_k), \text{ at } \partial \Omega_{\Gamma_1}, \qquad \nabla \Psi \cdot \nu = 0, \text{ at } \partial \Omega_N,$$

$$\Psi = 0, \quad \text{at } \partial \Omega_{\Gamma_0},$$

$$-\nabla \cdot (a_k(x) \nabla u_k) = 0, \text{ in } \Omega, \qquad u_k = U(x), \text{ at } \partial \Omega_D, \qquad \nabla u_k \cdot \nu = 0, \text{ at } \partial \Omega_N.$$

Example: Problem with exact data.

Evolution of both iteration error and residual is shown in Figure 5. The PLW method (GREEN) is compared with the LW method (BLUE) and with the SD method (RED).

Second example: Problem with noisy data.

Evolution of both iteration error and residual is shown in Figure 6. The PLW method (GREEN) is compared with the LW method (BLUE) and with the SD method (RED). The stop criteria (20) is reached after nine steps of the PLW iteration, 22 steps for the SD iteration, and 153 steps for the LW iteration.

6. Conclusions

In this work, we use the TCC to devise a family of relaxed projection Landweber methods for solving operator Equation (2). The distinctive features of this family of methods are:

- the basic method in this family (the PLW method) outperformed, in our preliminary numerical experiments, the classical Landweber method as well as the steepest descent method (with respect to both the computational cost and the number of iterations);
- the PLW method is convergent for the constant of the TCC in a range *twice as large* as the one required for the convergence of Landweber and other gradient type methods;
- for noisy data, the iteration of the PLW method progresses towards the solution set for residuals twice as small as the ones prescribed by the discrepancy principle for Landweber [2, Equation (11.10)] and steepest descent [24, Equation (2.4)] methods. This follows from the fact that the constant prescribed by the discrepancy principle for our method and for Landweber/steepest-descent are, respectively

$$\tau = \frac{1+\eta}{1-\eta} \quad \text{and} \quad \tau = 2\frac{1+\eta}{1-2\eta};$$

- the proposed family of projection-type methods encompasses, as particular cases, the Landweber method, the steepest descent method as well as the minimal error method; thus, providing an unified framework for their convergence analysis.

In our numerical experiments, the *residue* in the PLW method has very strong oscillations for noisy data (Figure 3) and for exact data (Figure 2). Since this method iterations' aims to reduce the iteration error, a non-monotone behavior of the residual is to be expected. In ill-posed problem error and residual are poor correlated, which may explain the large variations on the second one observed in our experiments with the PLW. Up to now it is not clear to us why this non-monotonicity happened to be oscillatory in our experiments.

Although projection type methods for solving systems of *linear* equations dates back to [25, 26], the use of these methods for ill-posed equations is more recent, see, e.g. [14].

A family of relaxed projection gradient-type methods for solving *linear* ill-posed operator equations was proposed in [27]. In this work, we extended to the non-linear case, under the TCC, the analysis of [27].

References

[1] Bakushinsky, A., Kokurin, M. (2004). *Iterative Methods for Approximate Solution of Inverse Problems*, Vol. 577 of Mathematics and Its Applications. Dordrecht: Springer.

[2] Engl, H., Hanke, M., Neubauer, A. (1996). *Regularization of Inverse Problems*. Dordrecht: Kluwer Academic Publishers.

[3] Hanke, M., Neubauer, A., Scherzer, O. (1995). A convergence analysis of Landweber iteration for nonlinear ill-posed problems. *Numer. Math.* 72(1):21–37.

[4] Kaltenbacher, B., Neubauer, A., Scherzer, O. (2008). *Iterative Regularization Methods for Nonlinear Ill-Posed Problems*, Vol. 6 of Radon Series on Computational and Applied Mathematics. Berlin: Walter de Gruyter GmbH & Co. KG.

[5] Landweber, L. (1951). An iteration formula for Fredholm integral equations of the first kind. *Amer. J. Math.* 73(3):615–624.

[6] Morozov, V. (1993). *Regularization Methods for Ill–Posed Problems*. Boca Raton, FL: CRC Press.

[7] Scherzer, O. (1993). Convergence rates of iterated Tikhonov regularized solutions of nonlinear ill-posed problems. *Numer. Math.* 66(1):259–279.

[8] Seidman, T., Vogel, C. (1989). Well posedness and convergence of some regularisation methods for non–linear ill posed problems. *Inverse Probl.* 5(2):227–238.

[9] Tikhonov, A. (1963). Regularization of incorrectly posed problems. *Soviet Math. Dokl.* 4:1624–1627.

[10] Tikhonov, A., Arsenin, V. (1977). *Solutions of Ill-Posed Problems*. Washington, D.C.: John Wiley & Sons, Translation editor: Fritz John.

[11] Herman, G. T. (1975). A relaxation method for reconstructing objects from noisy X-rays. *Math. Program.* 8(1):1–19.

[12] Herman, G. T. (1980). *Image Reconstruction from Projections*. New York–London: Academic Press, Inc. [Harcourt Brace Jovanovich, Publishers].

[13]　Natterer, F. (1977). Regularisierung schlecht gestellter probleme durch projektions-verfahren. *Numer. Math.* 28(3):329–341.

[14]　Natterer, F. (1986). *The Mathematics of Computerized Tomography*. B. G. Teubner, Stuttgart, Chichester: John Wiley & Sons, Ltd.

[15]　Eicke, B. (1992). Iteration methods for convexly constrained ill-posed problems in Hilbert space. *Numer. Punct. Anal. Optim.* 13(5–6):413–429.

[16]　Neubauer, A., Scherzer, O. (1995). A convergence rate result for a steepest descent method and a minimal error method for the solution of nonlinear ill-posed problems. *Z. Anal. Anwend.* 14(2):369–377.

[17]　Vasin, V. V., Eremin, I. I. (2009). *Operators and Iterative Processes of Fejér Type. Inverse and Ill-Posed Problems Series*. Berlin: Walter de Gruyter GmbH & Co. KG.

[18]　Burger, M., Engl, H. W., Leitão, A., Markowich, P. (2004). On inverse problems for semiconductor equations. *Milan J. Math.* 72(1):273–313.

[19]　Leitão, A. (2006). Semiconductors and Dirichlet-to-Neumann maps. *Mat. Appl. Comput.* 25(2–3):187–203.

[20]　Leitao, A., Markowich, P., Zubelli, J. (2006). Inverse problems for semiconductors: models and methods. In: Cercignani, C., Gabetta, E., eds. *Chapter in Transport Phenomena and Kinetic Theory: Applications to Gases, Semiconductors, Photons, and Biological Systems*. Boston, MA: Birkhäuser, pp. 117–149.

[21]　Leitao, A., Markowich, P., Zubelli, J. (2006). On inverse doping profile problems for the stationary voltage-current map. *Inv. Probl.* 22(3):1071–1088.

[22]　Lechleiter, A., Rieder, A. (2008). Newton regularizations for impedance tomography: convergence by local injectivity. *Inv. Probl.* 24(6):065009–065018.

[23]　Bank, R. E. (1994). *PLTMG: A Software Package for Solving Elliptic Partial Differential Equations, Volume 15 of Frontiers in Applied Mathematics*. Users' guide 7.0. Philadelphia, PA: Society for Industrial and Applied Mathematics (SIAM).

[24]　Scherzer, O. (1996). A convergence analysis of a method of steepest descent and a two-step algorithm for nonlinear ill-posed problems. *Numer. Funct. Anal. Optim.* 17(1–2):197–214.

[25]　Cimmino, G. (1938). Calcolo approssimato per le soluzioni dei sistemi di equazioni lineari. *La Ricerca Scientifica*. XVI(9):326–333.

[26]　Kaczmarz, S. (1937). Angenäherte auflösung von systemen linearer gleichungen. *Bull. Int. Acad. Pol. Sci. Let. Cl. Med.* 35:355–357.

[27]　McCormick, S., Rodrigue, G. (1975). A uniform approach to gradient methods for linear operator equations. *J. Math. Anal. Appl.* 49(2):275–285.