Projective acceleration of the nonlinear Landweber method under the tangential cone condition

A. Leitão[†] B. F. Svaiter[‡]

May 13, 2015

Abstract

We use convex analysis to devise an iterative method for solving ill-posed problems modeled by nonlinear operators acting between Hilbert spaces. The *tangential cone condition* (TCC) is our starting point for defining special convex sets possessing a separation property. This is the key ingredient to define an iterative method based on successive orthogonal projections onto these sets. Since the stepsize of this new method happens to be a multiple of the Landweber (LW) iteration stepsize, we call this new method *projective LW* (PLW). Conversely, by introducing relaxation in the projection of the PLW method we obtain a family of methods that include LW and LW method with line-search (LWls). Moreover, the convergence analysis of this family provides a unified framework for all above mentioned methods under the TCC. Numerical experiments are presented for a nonlinear 2D elliptic parameter identification problem, validating the efficiency of the PLW method compared with the well known LW and LWls iterations.

Keywords. Ill-posed problems; Nonlinear equations; Landweber method, Projective method.

AMS Classification: 65J20, 47J06.

1 Introduction

In this article we propose a new projective type method for obtaining stable approximate solutions of 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

[†]Department of Mathematics, Federal Univ. of St. Catarina, P.O. Box 476, 88040-900 Florianópolis, Brazil. acgleitao@gmail.com.

[‡]IMPA, Estr. Dona Castorina 110, 22460-320 Rio de Janeiro, Brazil. benar@impa.br.

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 is not know. Instead, what is available is only approximate measured data $y^\delta \in Y$ satisfying

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

where $\delta > 0$ is the noise level (assumed here to be known). Thus, the abstract formulation of the inverse problems under consideration is to find x such that

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

Standard methods for obtaining stable solutions of the operator equation in (2) can be divided in two major groups, namely, *Iterative type* regularization methods [1, 6, 9, 13, 14] and *Tikhonov type* regularization methods [6, 19, 23, 24, 25, 22]. A classical and general condition commonly used in the convergence analysis of these methods is the *Tangent Cone Condition* (TCC) [9].

In this work we use the TCC to devise a new projective method for solving (2). The distinctive features of this method are as follows:

- it outperformed, in our preliminary numerical experiments, the classical Landweber iteration as well as its line-search variant (with respect to both the computational cost and the number of iterations);
- by introducing relaxation in its stepsize, we obtain a family of projective-type methods which encompasses, as particular cases, the two above mentioned methods; thus, providing an unified framework for their convergence analysis.

In view of theses features, this new method will be called Projected Landweber (PLW) method.

Next we review the Landweber and briefly present the PLW method for solving the operator equation (2).

The Landweber (LW) method: If the operator F is Fréchet differentiable at $D(F) \subset X$, the Landweber method for solving (2) is defined by

$$x_{k+1}^\delta \;:=\; x_k^\delta \;-\; F'(x_k^\delta)^*\big(F(x_k^\delta) - y^\delta\big)\,. \tag{3}$$

Here $F'(z) : X \to Y$ is the Fréchet derivative of F at $z \in D(F)$ and $F'(z)^* : Y \to X$ is the corresponding adjoint operator. A convergence analysis for this method (including convergence rates) can be found, e.g., in [6] (see also [13] and the references therein). This method was originally proposed by L. Landweber [14], who was interested in proving strong convergence of the method of successive approximations for the fixed point equation

$x = x - F^*(Fx - y)$, which relates to the normal equation $F^*Fx = F^*y$ in the particular case when $F \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$ in (2) is a linear compact operator.

Under appropriate assumptions on the initial guess x_0^δ and on the nonlinear operator F [6, Sec. 11.1], in particular the TCC (see (6) below) and $\|F'(x)\| \leq 1$ in a suitable ball, one can prove: monotonicity of the iteration error; square summability of the residual series; strong convergence of the sequence x_k (for exact data $y^\delta = y$) towards some $x^* \in D(F)$ with $F(x^*) = y$; semi-convergence of x_k^δ (in the case of noisy data). Moreover, with additional assumptions on the solution set (the so called source conditions [13, Eq. (2.22)]), rates of convergence for the iteration error $\|x_k^\delta - x^*\|$, as well as for the residual $\|F(x_k^\delta) - y^\delta\|$ can be obtained.

The projective Landweber (PLW) method: The starting point for the derivation of the PLW method proposed in this manuscript is the above mentioned TCC.

Given a point $x \in D(F)$, we use the TCC to construct a half-space H_x (7) which separates x from the solution set $F^{-1}(y)$, i.e., $F^{-1}(y) \subset H_x$ and $x \notin H_x$ (see Lemma 3.1). Consequently, the orthogonal projection of x onto H_x is a better approximation to $F^{-1}(y)$ than the point x itself. The new iterate is constructed orthogonally projecting x onto this separating set H_x. Therefore, the PLW method is derived as a successive projection method which is strongly related with the Landweber method, because $F'(x)^*(F(x) - y)$ is normal to the separating half-space H_x

To further accelerate the method and make it more general, we consider under/over relaxed projections. The resulting iterative method can be written in the form

$$x_{k+1}^\delta \;:=\; x_k^\delta \;- \theta_k \lambda_k \; F'(x_k^\delta)^* \big(F(x_k^\delta) - y^\delta\big)$$

where $\theta_k \in (0, 2)$ is a relaxation parameter and λ_k gives the exact projection of x_k^δ onto $H_{x_k^\delta}$ (for $\theta_k = 1$), see (9).

The use of projection methods for solving linear ill-posed problems dates back to the 70's (with the seminal works of Frank Natterer and Gabor Herman) [20, 21, 10, 11]. The *combination* of Landweber iterations with projections onto a feasible set, for solving (2) with y^δ in a convex set was analyzed in [].

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 PLW method. In Section 3 we define the convex sets H_x (7), prove a special separation property (Lemma 3.1) and introduce the PLW method (9). 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 stopping criteria (21), 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 elementary 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

$$F_\delta(x) \;:=\; F(x) - y^\delta \qquad \text{and} \qquad F_0(x) \;:=\; F(x) - y\,. \tag{4}$$

Throughout this work we make the following assumptions, which are frequently used in the analysis of iterative regularization methods (see, e.g., [6, 13, 22]):

A1 F is a continuous operator defined on $D(F) \subset X$, which has nonempty 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

$$\|F'(x)\| \;\leq\; C\,, \qquad x \in B_\rho(x_0) \subset D(F) \tag{5}$$

(the point x_0 will be used as initial guess for the PLW iteration).

A2 The *local tangential cone condition* (TCC) [6, 13]

$$\|F(\bar{x}) - F(x) - F'(x)(\bar{x} - x)\|_Y \;\leq\; \eta \|F(\bar{x}) - F(x)\|_Y\,, \quad \forall\, x, \bar{x} \in B_\rho(x_0) \tag{6}$$

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 [6]. The TCC (6) represents a uniform assumption (on a ball of radius ρ) on the non-linearity of the operator F, and has interesting consequences (see [6, pg.278–280] or [13, pg.6 and Sec.2.4 (pg.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*

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

Proof. Item 1 follows immediately from the TCC and the triangle inequality, as proved in [6, Eq.(11.7)].

Direct algebraic manipulations yield

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

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

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

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 $\|F_0(x)-F_0(\bar{x})\| \leq \|F_0(x)\| + \|F_o(\bar{x})\|$. □

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) $\lim_{k\to\infty} \|F'(x_k)^* F_0(x_k)\| = 0$; *b)* $\lim_{k\to\infty} \|F_0(x_k)\| = 0$; *c)* $F(\bar{x}) = y$.

Proof. See [13, pg. 279] for a proof of the first assertion. For proving the second assertion:
the implication $(b) \Rightarrow (a)$ follows from **A1** and the hypothesis $(x_k) \in B_\rho(x_0)$;
$(a) \Rightarrow (b)$ follows from Proposition 2.1, item 3 with $x = x_k$ and $\bar{x} = x^\star$;
$(b) \Rightarrow (c)$ and $(b) \Leftarrow (c)$ follow from the hypothesis $\lim_{k\to\infty} \|x_k - \bar{x}\| = 0$ and **A1**. □

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 (6) for $\bar{x} = x^\star \in F^{-1}(y)$ using notation (4).

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 The projective Landweber (PLW) method

In this section we assume that exact data $y^\delta = y \in \mathrm{Rg}(F)$ are available, introduce the PLW method and its relaxed variants for the exact data case, and prove their convergence.

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

$$H_x \;:=\; \{z \in X \mid \langle z - x, F'(x)^* F_0(x)\rangle \leq -(1-\eta)\|F_0(x)\|^2\}\,. \tag{7}$$

Note that H_x is either $\emptyset$, a closed half-space, or X. As we will prove next, the set 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*

$$0 \;\geq\; (1-\eta)\,\|F_0(x)\|^2 + \langle F'(x)^* F_0(x), x^\star - x\rangle\,, \quad \forall x^\star \in B_\rho(x_0) \cap F^{-1}(y)\,. \tag{8}$$

Consequently,

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

Proof. The first statement follows trivially from Proposition 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 (7). □

We are now ready to introduce the PLW method and its relaxed variants. Choose $x_0 \in X$ according to **A2** and **A3** and define, for $k \geq 0$, the sequence

$$x_{k+1} := x_k - \theta_k\,\lambda_k\,F'(x_k)^* F_0(x_k), \tag{9a}$$

$$\text{where } \theta_k \in (0,2),\; \lambda_k := \begin{cases} 0, & \text{if } F'(x_k)^* F_0(x_k) = 0 \\ \dfrac{(1-\eta)\|F_0(x_k)\|^2}{\|F'(x_k)^* F_0(x_k)\|^2}, & \text{otherwise.} \end{cases} \tag{9b}$$

In view of definition (7), 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 (9) 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 *relaxed PLW methods* are obtained by choosing $\theta_k \in (0, 2)$, which amounts to taking x_{k+1} as a relaxed orthogonal projection of x_k onto H_{x_k}.

Iteration (9) 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 relaxed PLW methods in the case of exact data, i.e., $\|x^\star - x_{k+1}\| \geq \|x^\star - x_k\|$, whenever $\theta_k \in (0, 2)$.

Proposition 3.2. *Let assumptions **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 (9), then*

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

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 (9) and elementary properties of over/under relaxed orthogonal projections. □

Direct inspection of the inequality in Proposition 3.2 shows that the choice $\theta_k \in (0, 2)$, as prescribed in (9), 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 relaxed PLW iteration (for exact data) to *some* point $\bar{x} \in B_\rho(x_0)$. Theorem 3.4 gives a sufficient condition for strong convergence of the relaxed PLW iteration to *a solution* of $F(x) = y$, shows that Landweber method with line search as well as Landweber method are particular instances of relaxed PLW methods, and proves convergence of these two methods within this framework.

Recall that the Landweber iteration with line search is given by

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

Theorem 3.3. *If assumptions **A1** – **A3** hold true, then the sequences (x_k), (θ_k) as specified in (9) are well defined and*

$$x_k \in B_{\rho/2}(x^\star) \subset B_\rho(x_0)\,,\ \forall\ k \in \mathbb{N}\,. \tag{10}$$

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

$$(1-\eta)^2 \sum_{k=0}^{\infty} \theta_k \frac{\|F_0(x_k)\|^4}{\|F'(x_k)^* F_0(x_k)\|^2} \ <\ \infty \tag{11}$$

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

Theorem 3.4. *Let assumptions **A1** – **A3** hold true, and the sequences (x_k), (θ_k) be defined as in (9). The following assertions 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 assumption **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 $0 < \theta_k \le (1-\eta)^{-1} < 2$, iteration (9) reduces to the Landweber method with line-search strategy (LWls) and (x_k) converges to some $\bar{x} \in B_\rho(x_0)$ solving $F(\bar{x}) = y$.

*c) If assumptions **A1** and **A2** hold with $C \le 1$ 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 $0 < \theta_k \le (1-\eta)^{-1} < 2$, iteration (9) reduces to the classical Landweber iteration ((3) with $\delta = 0$) 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 (10) holds for $k = 0$. Suppose that the sequence (x_k), is well defined up to k_0 and that (10) 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 (9) and Proposition 3.2 that (10) 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 (10), Proposition 2.2 and (9), that $x_j = x_{\tilde{k}}$ for $j \ge \tilde{k}$, 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 (10) and Proposition 2.2 that $F'(x_k)^* F_0(x_k) \neq 0$ for all k. According to (9b)

$$\lambda_k \;:=\; (1-\eta)\,\|F_0(x_k)\|^2\,\|F'(x_k)^* F_0(x_k)\|^{-2}. \tag{12}$$

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

$$(1-\eta)^2 \sum_{k=0}^{\infty} \theta_k \frac{\|F_0(x_k)\|^4}{\|F'(x_k)^* F_0(x_k)\|^2} \;=\; (1-\eta) \sum_{k=0}^{\infty} \theta_k \lambda_k \|F_0(x_k)\|^2 \;<\; \infty\,, \tag{13}$$

which, in particular, proves (11).

If $\sum \theta_k \lambda_k < \infty$ then $\sum \|x_k - x_{k+1}\| < \infty$ (due to (9a) and Assumption **A1**) and (x_k) is a Cauchy sequence.

Suppose that $\sum \theta_k \lambda_k = \infty$. It follows from (13) 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

$$0 \leq k \leq \ell_i \;\Rightarrow\; \|F_0(x_k)\| \;\geq\; \|F_0(x_{\ell_i})\|\,. \tag{14}$$

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

$$\begin{aligned}
\|x_k - z\|^2 &= \|x_{k+1} - z\|^2 - \|x_k - x_{k+1}\|^2 - 2\langle x_k - x_{k+1}, x_k - z\rangle \\
&\leq \|x_{k+1} - z\|^2 - 2\langle x_k - x_{k+1}, x_k - z\rangle \\
&= \|x_{k+1} - z\|^2 + 2\theta_k \lambda_k \langle F'(x_k)^* F_0(x_k), x_k - z\rangle \\
&\leq \|x_{k+1} - z\|^2 + 8\lambda_k (\|F_0(x_k)\|^2 + \|F_0(x_k)\|\|F_0(z)\|)\,,
\end{aligned} \tag{15}$$

where the second equality follows from (9a) and the last inequality follows from Proposition 2.1, item 2, and the assumption $\eta < 1$. Thus, taking $z = x_{\ell_i}$ in (15), 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 (13) that $\lim_{m\to\infty} s_m = 0$. If $0 \leq k < \ell_i$, by adding the above inequality for $j = k, k+1, \ldots, \ell_i - 1$, 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 Assumption **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 from Theorem 3.3, the definition of θ_k and from Assumption **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). □

Remark 3.5. *The argument used to establish strong convergence of sequence (x_k) in the proof of Theorem 3.3 is inspired in the technique used in [9, 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 [9], 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 (14) 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 present the PLW method and its relaxed variants for the noisy data case and investigate their 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 (4), that is, $F_\delta(x) = F(x) - y^\delta$.

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

$$H_x^\delta := \left\{ z \in X \;\middle|\; \langle z - x, F'(x)^* F_\delta(x) \rangle \leq -\|F_\delta(x)\| \Big((1-\eta)\,\|F_\delta(x)\| - (1+\eta)\delta \Big) \right\} . \tag{16}$$

Next we prove the "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*

$$0 \geq \|F_\delta(x)\| \big[(1-\eta)\,\|F_\delta(x)\| - (1+\eta)\,\delta \big] + \langle x^\star - x, F'(x)^* F_\delta(x) \rangle \,, \tag{17}$$

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 (6). 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 (17). □

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 that $F'(x)^* F_\delta(x) \neq 0$.

The iteration formula of the PLW method and its relaxed variants for inexact data is given by

$$x_{k+1}^\delta \;:=\; x_k^\delta - \theta_k \frac{p_\delta(\|F_\delta(x_k^\delta)\|)}{\|F'(x_k^\delta)^* F_\delta(x_k^\delta)\|^2} F'(x_k^\delta)^* F_\delta(x_k^\delta) \,,\; \theta_k \in (0,2) \,, \tag{18}$$

where

$$p_\delta(t) \;:=\; t((1-\eta)t - (1+\eta)\delta)\,, \tag{19}$$

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 $H^\delta_{x_k^\delta}$ and the relaxed variants, which use $\theta_k \in (0,2)$, define x_{k+1}^δ as a relaxed projection of x_k^δ onto $H^\delta_{x_k^\delta}$.

Let

$$\tau \;>\; \frac{1+\eta}{1-\eta}\,. \tag{20}$$

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

$$k_*^\delta \;:=\; \max\left\{k \in \mathbb{N}\,;\; \|F_\delta(x_j^\delta)\| > \tau\delta\,,\; j = 0,1,\ldots,k-1\right\}. \tag{21}$$

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 (18) is well defined for $k = 0,\ldots,k_*^\delta$.

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 $k = 0,\ldots,k_*^\delta$. 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 assumptions **A1** – **A3** hold true and θ_k be chosen as in (18). If $x_k^\delta \in B_\rho(x_0)$ and $\|F_\delta(x_k^\delta)\| > \tau\delta$, then*

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

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 $H^\delta_{x_k^\delta}$ with a relaxation factor θ_k. The conclusion follows from this fact, Lemma 4.1, the iteration formula (18), and elementary properties of over/under relaxed orthogonal projections. □

Theorem 4.3. *If Assumptions **A1** – **A3** hold true, then the sequences (x_k^δ), (θ_k) as specified in (18) (together with the stopping criterion (21)) are well defined and*

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

Moreover, if $\theta_k \in [a,b] \subset (0,2)$ for all $k \leq k_*^\delta$, then this stopping index k_*^δ defined in (21) 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 $k < k_*^\delta$

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

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

$$x_{x+1}^\delta \;=\; x_k^\delta - \; F'(x_k^\delta)^* F_\delta(x_k^\delta)$$

(which requires $\eta < 1/2$, $C \leq 1$ in **A1** – **A2**) using the classical discrepancy principle (21) with

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

corresponds to PLW for inexact data analyzed in Theorem 4.3 with

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

where the second inequality follows from Assumption **A3** and the third inequality follows from Lemma 4.1. Consequently, the convergence analysis for the PLW method for noisy data encompasses the Landweber iteration for noisy data (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 (the semi-convergence of the PLW method). Notice that this is the first time were the strong assumption **A4** is needed in the text.

Theorem 4.4. *Let assumptions **A1** – **A4** hold true. For each fixed $k \in \mathbb{N}$, the element x_k^δ, computed after kth-iterations of the PLW method (18), depends continuously on the data y^δ.*

Proof. From (20), assumptions **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 $x_k^\delta = \big(\varphi(\cdot, y^\delta, \delta)\big)^k(x_0)$ 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 (21) 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., [6, Theor. 11.5] or [13, Theor. 2.6]) and thus is omitted.

Theorem 4.5. *Let assumptions **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 (18) is stopped with $k_*^j := k_*^{\delta_j}$ according to the discrepancy (21), then $(x_{k_*^j}^\delta)$ converges strongly to a solution $\bar{x} \in B_\rho(x_0)$ of $F(x) = y$ as $j \to \infty$.*

5 Numerical experiments

In what follows 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 LWls.

5.1 Description of the mathematical model

We briefly introduce a model which plays a key rule in inverse doping problems with current flow measurements, namely the 2D *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 $U \equiv 0$ [16, 4], 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

[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 $U = 0$.

recombination-generation rate is present [18, 17]. 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

$$\Sigma_C'(0)h \;=\; \mu_n\, e^{V_{\rm bi}}\hat{u}_\nu - \mu_p\, e^{-V_{\rm bi}}\hat{v}_\nu \;\in\; H^{1/2}(\Gamma_1)\,, \tag{22}$$

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

$$\begin{aligned}
\operatorname{div}(\mu_n e^{V^0}\nabla\hat{u}) &= 0 && \text{in } \Omega && \text{(23a)}\\
\operatorname{div}(\mu_p e^{-V^0}\nabla\hat{v}) &= 0 && \text{in } \Omega && \text{(23b)}\\
\hat{u} &= -\hat{v} = -h && \text{on } \partial\Omega_D && \text{(23c)}\\
\nabla\hat{u}\cdot\nu &= \nabla\hat{v}\cdot\nu = 0 && \text{on } \partial\Omega_N && \text{(23d)}
\end{aligned}$$

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

$$\begin{aligned}
\lambda^2\,\Delta V^0 &= e^{V^0} - e^{-V^0} - C(x) && \text{in } \Omega && \text{(24a)}\\
V^0 &= V_{\rm bi}(x) && \text{on } \partial\Omega_D && \text{(24b)}\\
\nabla V^0\cdot\nu &= 0 && \text{on } \partial\Omega_N\,. && \text{(24c)}
\end{aligned}$$

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 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, $V_{\rm bi}$ is a given logarithmic function [4].

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 [17, 18].

5.2 The inverse doping problem

The inverse problem we are concerned with consists in determining the doping profile function C in (24) from measurements of the linearized VC-map

[2] These concentrations are here written in terms of the Slotboom variables [17].

$\Sigma'_C(0)$ in (22), under the assumption $\mu_p = 0$ (the so-called *linearized stationary unipolar model close to equilibrium*). Notice that we can split the inverse problem in two parts:

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

$$\operatorname{div}(\mu_n \gamma \nabla \hat{u}) = 0 \text{ in } \Omega \qquad \hat{u} = -U(x) \text{ on } \partial\Omega_D \qquad \nabla \hat{u} \cdot \nu = 0 \text{ on } \partial\Omega_N. \tag{25}$$

for γ, from measurements of $\big[\Sigma'_C(0)\big](U) = \big(\mu_n \gamma \hat{u}_\nu\big)|_{\Gamma_1}$.

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

Since the evaluation of C from γ can be explicitly performed in a stable way, we shall focus on the problem of identifying the function parameter γ in (25). Summarizing, the inverse doping profile problem in the linearized stationary unipolar model (close to equilibrium) reduces to the identification of the parameter function γ in (25) from measurements of the Dirichlet-to-Neumann map $\Lambda_\gamma : H^{1/2}(\partial\Omega_D) \ni U \mapsto \big(\mu_n \gamma \hat{u}_\nu\big)|_{\Gamma_1} \in H^{-1/2}(\Gamma_1)$.

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 γ has to be performed from a finite number of measurements, i.e. from the data $\big\{(U_i, \Lambda_\gamma(U_i))\big\}_{i=1}^N \in \big[H^{1/2}(\Gamma_0) \times H^{-1/2}(\Gamma_1)\big]^N$.

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

$$F(\gamma) = \Lambda_\gamma(U) =: y\,, \tag{26}$$

where U is a fixed voltage profile satisfying the above assumptions, $X := L^2(\Omega) \supset D(F) := \{\gamma \in L^\infty(\Omega);\ 0 < \gamma_m \leq \gamma(x) \leq \gamma_M, \text{ a.e. in } \Omega\}$ and $Y := H^{1/2}(\Gamma_1)$. The operator F above is known to be continuous [4].

5.3 First experiment: The Calderon setup

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

1. the map $\gamma \mapsto u$ (solution of (25)) satisfies the TCC with respect to the $H^1(\Omega)$ norm [13];

2. it was proven in [15] that the discretization of the operator F in (26) using the finite element method (and basis functions constructed by a Delaunay triangulation) satisfies the TCC (6).

Therefore, the analytical convergence results of the previous sections do apply to finite-element discretizations of (26) 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 $(0,1) \times (0,1)$ and the above mentioned boundary parts are $\Gamma_1 = \partial\Omega_D := \partial\Omega$, $\Gamma_0 = \partial\Omega_N := \emptyset$.

• The parameter space is $H^1(\Omega)$ and the function $\gamma^\star(x)$ to be identified is shown in Figure 1.

• The fixed Dirichlet input for the DtN map (25) 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 (25) are plotted).

• The TCC constant η in (6) is not known for this particular setup. In our computations we used the value $\eta = 0.45$ which is in agreement with assumption **A2**.
(Note that the convergence analysis of the PLW method requires $\eta < 1$ while the nonlinear LW method requires the TCC with $\eta < 0.5$ [13, assumption (2.4)]. The above choice allows the comparison of both methods.)

• The "exact data" y in (26) is obtained by solving the direct problem (25) using a finite element type method and adaptive mesh refinement (approx 131.000 elements). In order to avoid inverse crimes, a coarser grid (with approx 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 (21) we assumed exact knowledge of the noise level and chose $\tau = 3$ in (20), which is in agreement with the above choice for η.

Remark 5.1 (Choosing the initial guess)**.** *The initial guess γ_0 used for all iterative methods is presented in Figure 1. According to Assumptions **A1** – **A3**, γ_0 has to be sufficiently close to $\gamma^\star$ (otherwise the PLW method may not converge). With this in mind, we choose γ_0 as the solution the Dirichlet boundary value problem*

$$\Delta\gamma_0 \;=\; 0\,,\ \textit{in}\ \Omega\,, \qquad \gamma_0 \;=\; U(x)\,,\ \textit{at}\ \partial\Omega\,.$$

This choice is an educated guess that incorporate the available a priori *knowledge about the exact solution $\gamma^\star$, namely: $\gamma_0 \in H^1(\Omega)$ and $\gamma_0 = \gamma^\star$ at $\partial\Omega_D$. Moreover, $\gamma_0 = \arg\min\{\|\nabla\gamma\|^2_{L^2(\Omega)} \mid \gamma \in H^1(\Omega),\ \gamma_{\partial\Omega} = \gamma^\star_{\partial\Omega}\}$.*

Remark 5.2 (Computing the iterative step)**.** *The computation of the kth-step of the PLW method (see (9)) requires the evaluation of* $F'(\gamma_k)^* F_0(\gamma_k)$. *According to [4], for all test functions* $v \in H_0^1(\Omega)$ *it holds*

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

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

$$-\nabla\cdot(\gamma_k\,\nabla V) \;=\; \nabla\cdot(v\,\nabla F(\gamma_k))\,,\ \textit{in}\ \Omega\,, \qquad V \;=\; 0\,,\ \textit{at}\ \partial\Omega\,.$$

Furthermore, in [4] it is shown that for all $\psi \in L^2(\partial\Omega)$ *and* $v \in H_0^1(\Omega)$

$$\langle F'(\gamma_k)^\times \psi, v\rangle_{L^2(\Omega)} \;=\; \langle \psi, V\rangle_{L^2(\partial\Omega)} \;=\; \langle \nabla\Psi\cdot\nabla u_k, v\rangle_{L^2(\Omega)}\,, \tag{27}$$

where Ψ, $u_k \in H^1(\Omega)$ *solve*

$$-\nabla\cdot(\gamma_k\,\nabla\Psi) \;=\; 0\,,\ \textit{in}\ \Omega\,, \qquad \Psi \;=\; \psi\,,\ \textit{at}\ \partial\Omega \tag{28a}$$

$$-\nabla\cdot(\gamma_k\,\nabla u_k) \;=\; 0\,,\ \textit{in}\ \Omega\,, \qquad u_k \;=\; U(x)\,,\ \textit{at}\ \partial\Omega\,. \tag{28b}$$

respectively. An direct consequence of (27), (28) is the variational identity

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

where Ψ *solves (28a) with* $\psi = F_0(\gamma_k)$.

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

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

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*

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

Within this setting, the PLW iteration (9) becomes

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

The iterative steps of the benchmark iterations LW and LWls (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 (18)).*

For solving the elliptic PDE's above described, needed for the implementation of the iterative methods, we used the package PLTMG [2] compiled with GFORTRAN-4.8 in a INTEL(R) Xeon(R) CPU E5-1650 v3.
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 LW method with line-search (LWls, RED). For comparison purposes, if one decides to stop iterating when $\|F_0(\gamma_k)\| < 0.025$ is satisfied, the PLW method needs only 43 iterations, while the LWls 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 LW method with line-search (LWls, RED). The stop criteria (21) is reached after 10 steps of the PLW iteration, 13 steps for the LWls iteration, and 32 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 $\partial\Omega_D \subsetneq \partial\Omega$, and $\Gamma_0 \neq \emptyset$, $\partial\Omega_N \neq \emptyset$.

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 γ is performed from a single measurement. To the best of our knowledge, within this setting assumptions (A1) - (A3) where not yet established for the operator F in (26) and its discretizations. Therefore, although the operator F is continuous [4], it is still unclear whether the analytical convergence results of the previous sections hold here.

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 $\gamma^\star(x)$ to be identified;
  - — The computation of the "exact data" y in (26);
  - — The choice for the TCC constant η in (6) and for τ in (20);
  - — 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 (25) 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 (25) are plotted.

• The initial condition γ_0 used for all iterative methods is shown in Figure 4 (b) and is given by the solution of the mixed boundary value problem

$$\Delta\gamma_0 \;=\; 0\,,\ \text{in}\ \Omega\,, \qquad \gamma_0 \;=\; U(x)\,,\ \text{at}\ \partial\Omega_D\,, \qquad \nabla\gamma_0\cdot\nu \;=\; 0\,,\ \text{at}\ \partial\Omega_N\,,$$

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

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

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

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

and Ψ, u_k solve

$$\begin{aligned} &-\nabla\cdot(\gamma_k\,\nabla\Psi) = 0,\ \text{in}\ \Omega, \quad \Psi = F_0(\gamma_k),\ \text{at}\ \partial\Omega_{\Gamma_1}, \qquad \nabla\Psi\cdot\nu = 0,\ \text{at}\ \partial\Omega_N,\\ &\qquad\qquad\qquad\qquad\quad \Psi = 0, \qquad\ \text{at}\ \partial\Omega_{\Gamma_0},\\ &-\nabla\cdot(\gamma_k\,\nabla u_k) = 0,\ \text{in}\ \Omega, \quad u_k = U(x), \quad \text{at}\ \partial\Omega_D, \qquad \nabla u_k\cdot\nu = 0,\ \text{at}\ \partial\Omega_N. \end{aligned}$$

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 LWls 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 LWls method (RED). The stop criteria (21) is reached after 9 steps of the PLW iteration, 17 steps for the LWls iteration, and 153 steps for the LW iteration.

6 Conclusions

In this work we use the TCC to devise a new projective method for solving operator equation (2). The distinctive features of this method are:

- it outperformed, in our preliminary numerical experiments, the classical Landweber iteration as well as its line-search variant (with respect to both the computational cost and the number of iterations);

- by introducing relaxation in its stepsize, we obtain a family of projective-type methods which encompasses, as particular cases, the two above mentioned methods; thus, providing an unified framework for their convergence analysis.

The last result provides a new interpretation of the Landweber iteration, under the TCC, as a projective method.

Projective type methods for the solution of systems of linear ill-posed equations are well known [12, 21], and there has been recent developments for the acceleration of such methods [5, 3, 3, 8, 7]. The analysis presented here points to the use of these projective type techniques for the efficient solution of systems of nonlinear ill-posed equations.

Acknowledgments

A.L. acknowledges support from the Brazilian research agencies CAPES, CNPq (grant 309767/2013–0), and from the AvH Foundation.
The work of B.F.S. was partially supported by CNPq (grants 474996/2013-1, 302962/2011-5) and FAPERJ (grant E-26/102.940/2011).

References

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

[2] R. E. Bank. *PLTMG: a software package for solving elliptic partial differential equations*, volume 15 of *Frontiers in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1994. Users' guide 7.0.

[3] J. Baumeister, B. Kaltenbacher, and A. Leitão. On levenberg-marquardt-kaczmarz iterative methods for solving systems of nonlinear ill-posed equations. *Inverse Probl. Imaging*, 4(3):335–350, 2010.

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

[5] J. Eckstein and B. F. Svaiter. General projective splitting methods for sums of maximal monotone operators. *SIAM J. Control Optim.*, 48(2):787–811, 2009.

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

[7] M. Haltmeier, A. Leitão, and O. Scherzer. Kaczmarz methods for regularizing nonlinear ill-posed equations. I. convergence analysis. *Inverse Probl. Imaging*, 1(2):289–298, 2007.

[8] M. Haltmeier, A. Leitão, and E. Resmerita. On regularization methods of EM-Kaczmarz type. *Inverse Problems*, 25:075008, 2009.

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

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

[11] G. T. Herman. *Image reconstruction from projections.* Academic Press, Inc. [Harcourt Brace Jovanovich, Publishers], New York-London, 1980. The fundamentals of computerized tomography, Computer Science and Applied Mathematics.

[12] S. Kaczmarz. Approximate solution of systems of linear equations. *Internat. J. Control*, 57(6):1269–1271, 1993.

[13] B. Kaltenbacher, A. Neubauer, and O. Scherzer. *Iterative regularization methods for nonlinear ill-posed problems*, volume 6 of *Radon Series on Computational and Applied Mathematics.* Walter de Gruyter GmbH & Co. KG, Berlin, 2008.

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

[15] A. Lechleiter and A. Rieder. Newton regularizations for impedance tomography: convergence by local injectivity. *Inverse Problems*, 24(6):065009, 18, 2008.

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

[17] A. Leitao, P. Markowich, and J. Zubelli. *Inverse Problems for Semiconductors: Models and Methods*, chapter in Transport Phenomena and Kinetic Theory: Applications to Gases, Semiconductors, Photons, and Biological Systems, Ed. C.Cercignani and E.Gabetta. Birkhäuser, Boston, 2006.

[18] A. Leitao, P. Markowich, and J. Zubelli. On inverse dopping profile problems for the stationary voltage-current map. *Inv.Probl.*, 22:1071–1088, 2006.

[19] V. Morozov. *Regularization Methods for Ill–Posed Problems.* CRC Press, Boca Raton, 1993.

[20] F. Natterer. Regularisierung schlecht gestellter Probleme durch Projektionsverfahren. *Numer. Math.*, 28(3):329–341, 1977.

[21] F. Natterer. *The mathematics of computerized tomography.* B. G. Teubner, Stuttgart; John Wiley & Sons, Ltd., Chichester, 1986.

[22] O. Scherzer. Convergence rates of iterated Tikhonov regularized solutions of nonlinear ill-posed problems. *Numer. Math.*, 66(2):259–279, 1993.

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

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

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

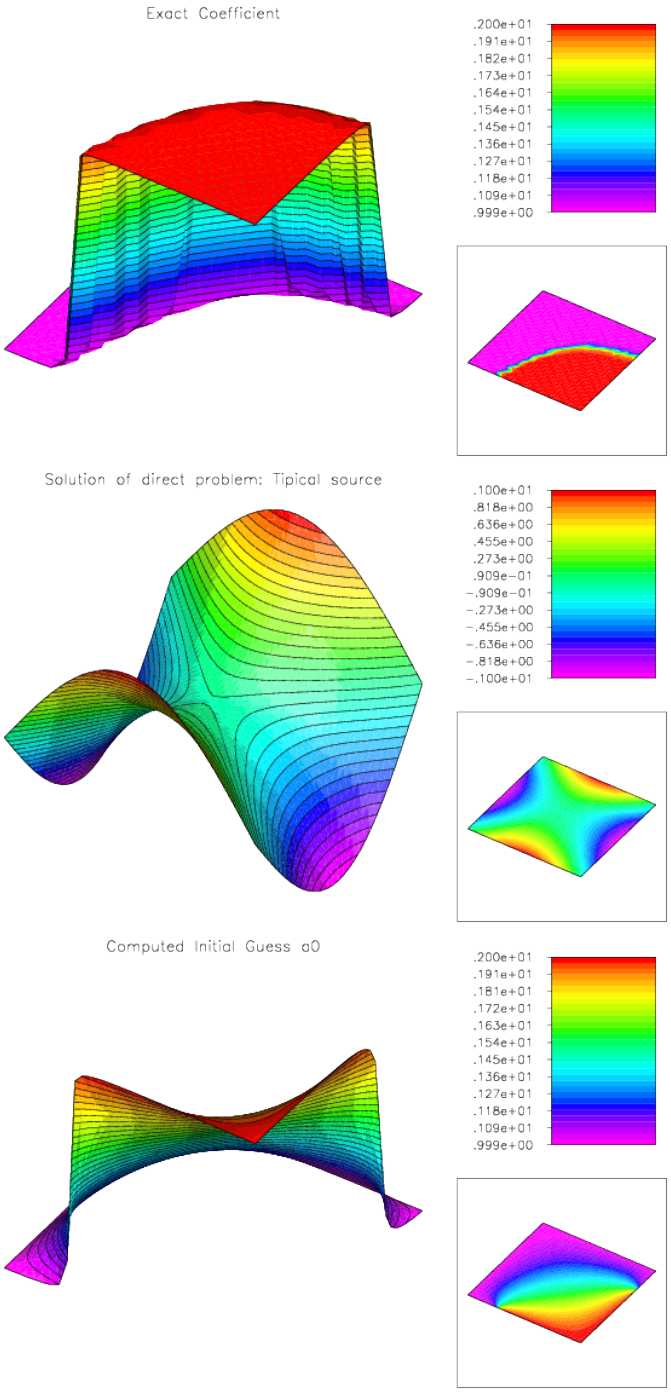

Figure 1: First experiment: setup of the problem. **Top:** The parameter $\gamma^*(x)$ to be identified; **Center:** Voltage source $U(x)$ – Dirichlet boundary condition at $\partial\Omega$ for the DtN map – and the corresponding solution $\hat{u}$ of (25); **Bottom:** Initial guess γ_0 for the iterative methods PLW, LW and LWls.

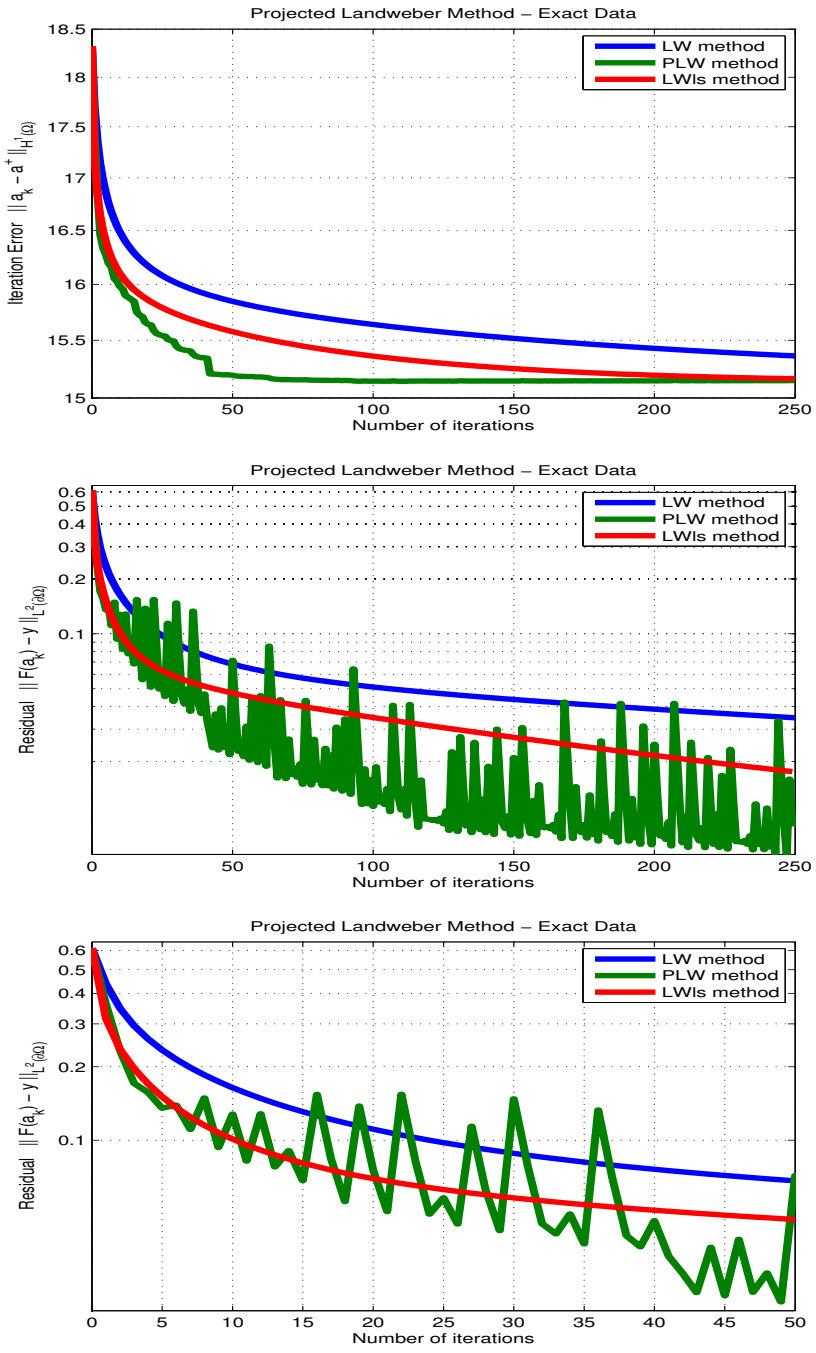

Figure 2: First experiment; example with exact data. The PLW method (GREEN) is compared with the LW method (BLUE) and with the LWls method (RED). **Top:** Evolution of the iteration error $\|\gamma_k - \gamma^\star\|_{H^1(\Omega)}$. **Middle:** Evolution of the residual $\|F(\gamma_k) - y\|_{L^2(\partial\Omega)}$. **Bottom:** Evolution of the residual, detail of the first 50 iterations.

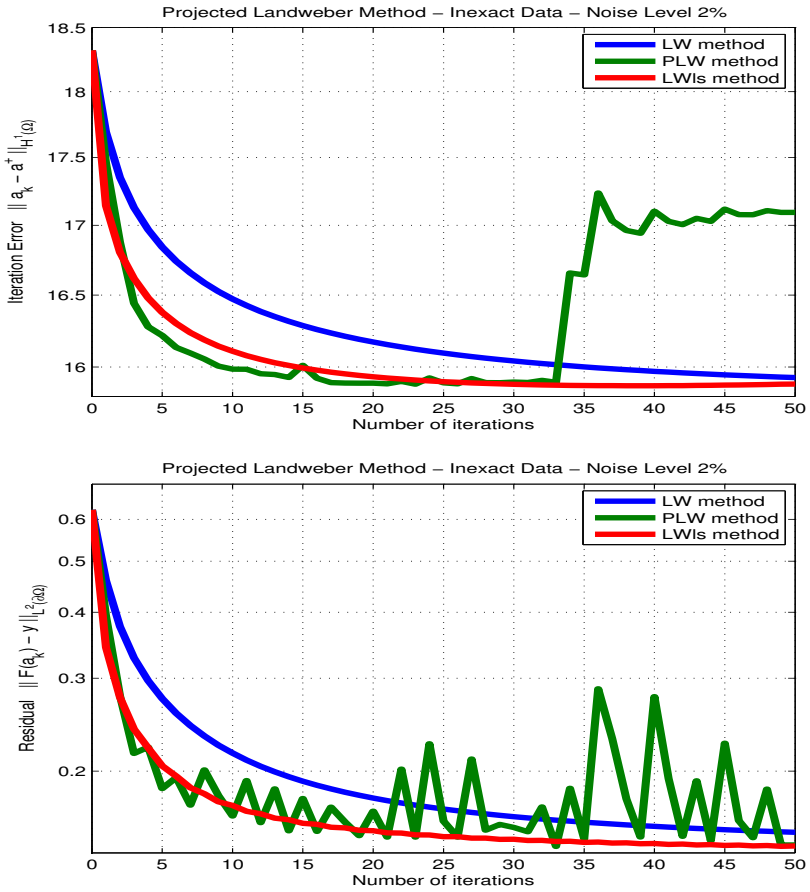

Figure 3: First experiment; example with noisy data. The PLW method (GREEN) is compared with the LW method (BLUE) and with the LWls method (RED). **Top:** Evolution of the iteration error $\|\gamma_k^\delta - \gamma^\star\|_{H^1(\Omega)}$. **Bottom:** Evolution of the residual $\|F(\gamma_k^\delta) - y^\delta\|_{L^2(\partial\Omega)}$.

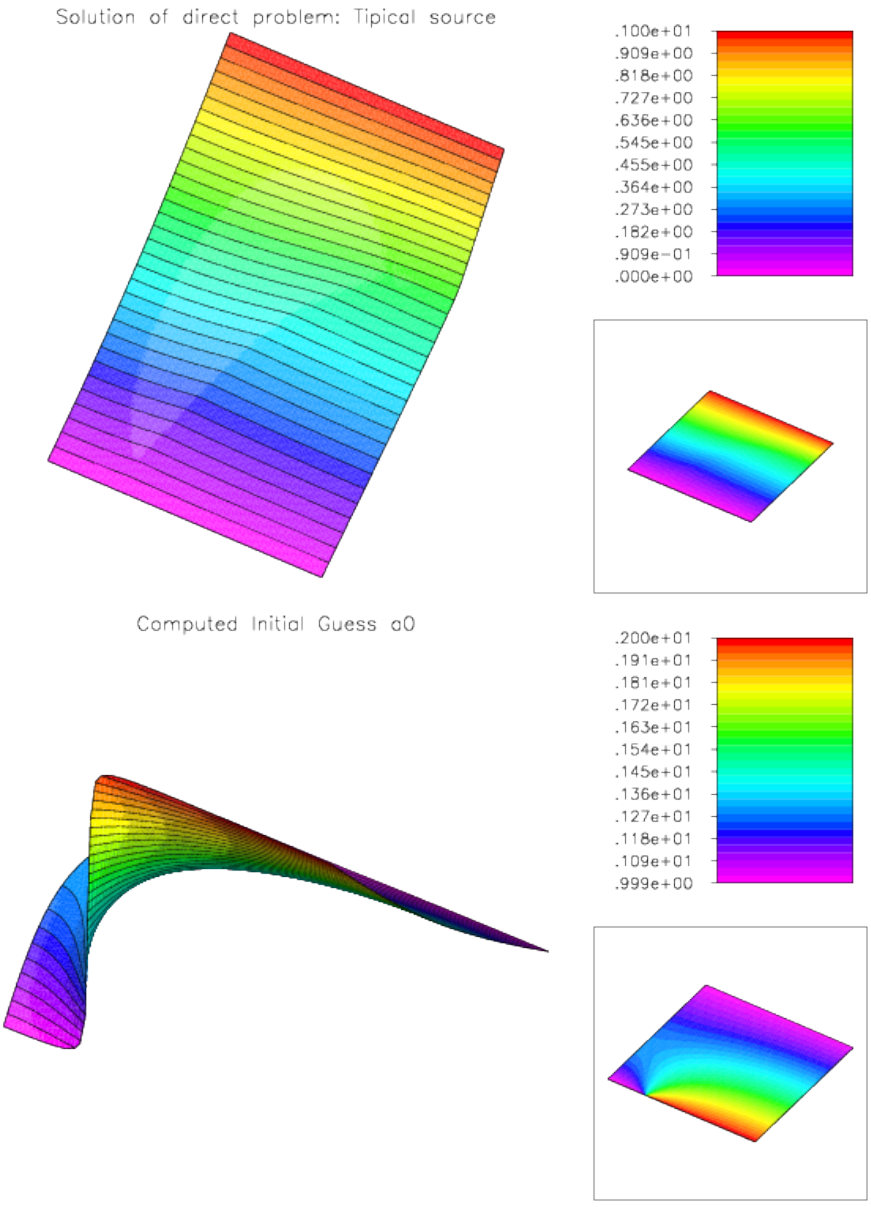

Figure 4: Second experiment: setup of the problem. **Top:** Voltage source $U(x)$ (Dirichlet boundary condition at $\partial\Omega_D$) for the DtN map and the corresponding solution $\hat{u}$ of (25); **Bottom:** Initial guess $\gamma_0 \in H^1(\Omega)$ satisfying $\gamma_0 = U$ at $\partial\Omega_D$ and $\nabla\gamma_0 \cdot \nu = 0$ at $\partial\Omega_N$.

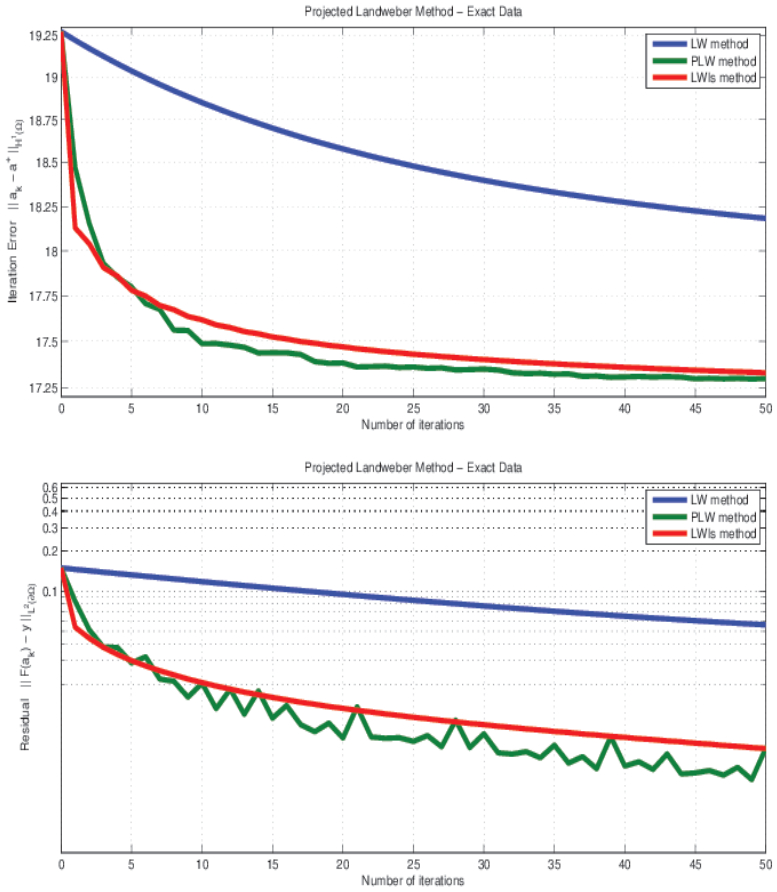

Figure 5: Second experiment: example with exact data. The PLW method (GREEN) is compared with the LW method (BLUE) and with the LWls method (RED). **Top:** Evolution of the iteration error $\|\gamma_k - \gamma^\star\|_{H^1(\Omega)}$. **Bottom:** Evolution of the residual $\|F(\gamma_k) - y\|_{L^2(\Gamma_1)}$.

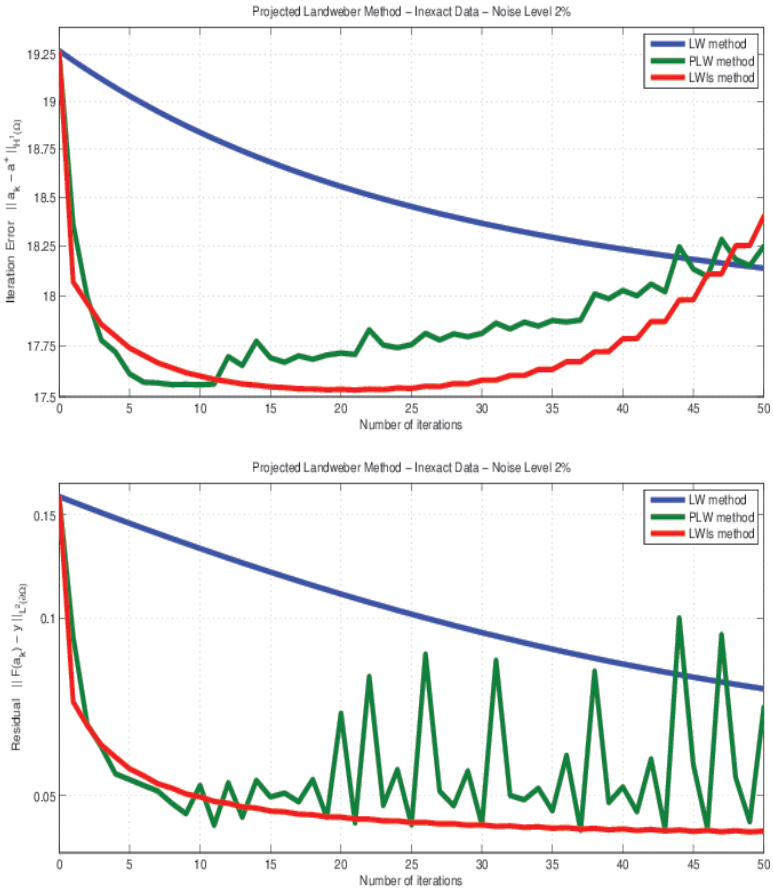

Figure 6: Second experiment: example with exact data. The PLW method (GREEN) is compared with the LW method (BLUE) and with the LWls method (RED). **Top:** Evolution of the iteration error $\|\gamma_k - \gamma^\star\|_{H^1(\Omega)}$. **Bottom:** Evolution of the residual $\|F(\gamma_k) - y\|_{L^2(\Gamma_1)}$.