Research Article

Rafaela Filippozzi, Eduardo Hafemann, Joel C. Rabelo, Fabio Margotti and Antonio Leitão*

A range-relaxed criteria for choosing the Lagrange multipliers in the Levenberg–Marquardt–Kaczmarz method for solving systems of non-linear ill-posed equations: Application to EIT-CEM with real data

https://doi.org/10.1515/jiip-2022-0007
Received January 24, 2022; accepted March 31, 2022

Abstract: In this article we propose and analyze a *Levenberg–Marquardt–Kaczmarz*-type (LMK) method for obtaining stable approximate solutions to systems of ill-posed equations modeled by non-linear operators acting between Hilbert spaces. We extend to the LMK iteration the strategy proposed in [A. Leitão, F. Margotti and B. F. Svaiter, Range-relaxed criteria for choosing the Lagrange multipliers in the Levenberg–Marquardt method, *IMA J. Numer. Anal.* **41** (2021), no. 4, 2962–2989] for choosing the Lagrange multipliers in the *Levenberg–Marquardt* (LM) method. Our main goal is to devise a simple (and easy to implement) strategy for computing the multiplier in each iterative step, such that the resulting LMK iteration is both stable and numerically efficient. Convergence analysis for the proposed LMK type method is provided, including convergence for exact data, stability and semi-convergence. Numerical experiments using real data are presented for a 2D parameter identification problem, namely the Electrical Impedance Tomography (EIT) problem. The mathematical model known as complete electrode model (EIT-CEM) is considered. The obtained numerical results validate the efficiency of the proposed LMK-type method.

Keywords: Ill-posed problems, non-linear equations, Levenberg–Marquardt method, Kaczmarz method

MSC 2010: 65J20, 47J06

1 Introduction

In this article we propose a *Levenberg–Marquardt–Kaczmarz*-type (LMK) method for regularizing systems of non-linear ill-posed operator equations. This is a Kaczmarz-type method [14], where each step is defined as

***Corresponding author: Antonio Leitão,** Department of Mathematics, Federal University of St. Catarina, 88040-900 Florianópolis SC, Brazil, e-mail: acgleitao@gmail.com
Rafaela Filippozzi, Fabio Margotti, Department of Mathematics, Federal University of St. Catarina, 88040-900 Florianópolis SC, Brazil, e-mail: rafaela.filippozzi@gmail.com, fabio.margotti@ufsc.br
Eduardo Hafemann, Department of Mathematics; and Department of Chemical Engineering, Federal University of St. Catarina, 88040-900 Florianópolis SC, Brazil, e-mail: hafemann.eduardo@gmail.com
Joel C. Rabelo, Department of Mathematics, Federal University of St. Catarina, 88040-900 Florianópolis SC; and Department of Mathematics, Federal University of Piaui, 64049-550, Teresina PI, Brazil, e-mail: joelrabelo@ufpi.edu.br

in the Levenberg–Marquardt method [11, 19, 21]. Additionally, the Lagrange multipliers are chosen as to guarantee the (linearized) residual of the next iterate to be in a *range* [18].

The *inverse problem* we are interested in consists of determining an unknown quantity $x \in X$ from the set of data $(y_0, \dots, y_{N-1}) \in Y^N$, where X and Y are Hilbert spaces, and $N \geq 1$. In practical situations, one does not know the data exactly. Instead, only approximate measured data $y_i^\delta \in Y$ satisfying

$$\|y_i^\delta - y_i\| \leq \delta_i, \quad i = 0, \dots, N-1, \tag{1.1}$$

are available, where $\delta_i > 0$ are the (known) noise levels. The available data y_i^δ are obtained by indirect measurements of the parameter x, this process being described by the system of ill-posed operator equations

$$F_i(x) = y_i, \quad i = 0, \dots, N-1, \tag{1.2}$$

where $F_i : X \to Y$ are given operators (see Section 2 for details).

Levenberg–Marquardt-type methods

Standard Levenberg–Marquardt (LM)-type methods for solving the ill-posed problem (1.1)–(1.2) are defined, after rewriting (1.2) as a single equation

$$\mathbf{F}(x) = \mathbf{y},$$

where $\mathbf{F} = (F_0, \dots, F_{N-1}) : X \to Y^N$ and $\mathbf{y} = (y_0, \dots, y_{N-1}) \in Y^N$, by the iteration formula

$$x_{k+1}^\delta = \arg\min_{x \in X} \{\|\mathbf{y}^\delta - \mathbf{F}(x_k^\delta) - \mathbf{F}'(x_k^\delta)(x - x_k^\delta)\|^2 + \alpha_k \|x - x_k^\delta\|^2\} \tag{1.3}$$

(see [2, 11]) or, equivalently, by

$$x_{k+1}^\delta = x_k^\delta + h_k^\delta, \text{ with } h_k^\delta := (\mathbf{F}'(x_k^\delta)^* \mathbf{F}'(x_k^\delta) + \alpha_k I)^{-1} \mathbf{F}'(x_k^\delta)^* (\mathbf{y}^\delta - \mathbf{F}(x_k^\delta)). \tag{1.4}$$

where $\mathbf{F}'(x_k^\delta) : X \to Y^N$ is the Fréchet derivative of $\mathbf{F}(x_k^\delta)$ and $\mathbf{F}'(x_k^\delta)^* : Y^N \to X$ is the adjoint operator to $\mathbf{F}'(x_k^\delta)$. The parameter $\alpha_k > 0$ can be viewed as the Lagrange multiplier of the problem of projecting x_k^δ onto a level-set of $\|\mathbf{y}^\delta - \mathbf{F}(x_k^\delta) - \mathbf{F}'(x_k^\delta)(x - x_k^\delta)\|^2$. If the sequence $\{\alpha_k = \alpha\}$ is constant, iteration (1.4) is called *stationary* LM [2], otherwise it is denominated *non-stationary* LM [6, 11, 18].

In the non-stationary LM methods, each α_k is chosen either *a priori* [6, 11] (e.g., the geometrical choice $\alpha_k = r^k$, $r < 1$) or *a posteriori* [11, 18]. In this article we focus on the *a posteriori* strategy investigated in [18], where the authors propose a choice for the Lagrange multipliers, which requires the linearized residual at the next iterate to assume a prescribed value dependent on the current residual and also on the noise level. More precisely, α_k is chosen so that x_{k+1}^δ satisfies

$$\Phi_1(\|\mathbf{F}(x_k^\delta) - \mathbf{y}^\delta\|, \delta) \leq \|\mathbf{y}^\delta - \mathbf{F}(x_k^\delta) - \mathbf{F}'(x_k^\delta)(x_{k+1}^\delta - x_k^\delta)\| \leq \Phi_2(\|\mathbf{F}(x_k^\delta) - \mathbf{y}^\delta\|, \delta),$$

with appropriately chosen functions Φ_i, $i = 1, 2$.

The LM-type methods may become inefficient if N is large or the evaluation of the step in (1.4) is expensive. In such cases, Kaczmarz-type methods which cyclically consider each equation in (1.2) separately, are reported to be faster [22] and are often the method of choice in practice. On the other hand, only few theoretical results about regularizing properties of LM-Kaczmarz methods are available, so far (see, e.g., [2, 3]).

Levenberg–Marquardt–Kaczmarz-type methods

The method proposed and analyzed in this manuscript for solving the ill-posed problem (1.1)–(1.2) is a Kaczmarz-type method, where each step is defined as in the LM method. Moreover, the choice of Lagrange multipliers is inspired in the one proposed in [18]. This iterative method can be summarized as follows:

$$x_{k+1}^\delta = x_k^\delta + h_k^\delta, \tag{1.5}$$

where

$$h_k^\delta = \begin{cases} (F'_{[k]}(x_k^\delta)^* F'_{[k]}(x_k^\delta) + \alpha_k I)^{-1} F'_{[k]}(x_k^\delta)^* (y_{[k]}^\delta - F_{[k]}(x_k^\delta)) & \text{if } \|F_{[k]}(x_k^\delta) - y_{[k]}^\delta\| > \tau\delta_{[k]}, \\ 0 & \text{otherwise,} \end{cases} \tag{1.6}$$

and

$$\alpha_k = \begin{cases} \text{chosen as in Algorithm 1} & \text{if } \|F_{[k]}(x_k^\delta) - y_{[k]}^\delta\| > \tau\delta_{[k]}, \\ \text{not defined} & \text{otherwise.} \end{cases} \tag{1.7}$$

Here $[k] = (k \bmod N) \in \{0, 1, \dots, N-1\}$, $x_0^\delta = x_0 \in X$ is an initial guess and $\tau > 1$ is a fixed constant (see next section).

The $h_k^\delta \in X$ in (1.6) is inspired in the iterative step proposed in [18] for the $[k]$th equation $F_{[k]}(x) = y_{[k]}$ of system (1.2), with data $y_{[k]}^\delta$ given as in (1.1). Notice that, if $\|F_{[k]}(x_k^\delta) - y_{[k]}^\delta\| \le \tau\delta_{[k]}$ for some k, then the computation of (α_k, h_k^δ) is avoided; we set $h_k^\delta = 0$ and $x_{k+1}^\delta = x_k^\delta$.

Following [18] we refer to this method as the *range-relaxed Levenberg–Marquardt–Kaczmarz* (rrLMK) method. Essentially, it consists in incorporating the Kaczmarz strategy to the Levenberg–Marquardt-type method in [18]. This procedure is analog to the one introduced in [8, 10], [5], [9], [3], and [7] regarding the Landweber Kaczmarz (LWK), the Steepest Descent Kaczmarz (SDK), the Expectation Maximization Kaczmarz (EMK), the iteratively regularized Gauss–Newton–Kaczmarz (irGNK), and the iterated Tikhonov Kaczmarz (iTK) iterations, respectively. It is worth mentioning that iteration (1.5), (1.6), (1.7) was considered in [2] for the constant choice $\alpha_k = \alpha$.

In Kaczmarz-type algorithms, a group of N subsequent steps (starting at some multiple of N) is called a cycle. The iteration (1.5)–(1.7) should be terminated when, for the first time, all x_k^δ are equal within a cycle. That is, we stop the iteration at step $k_* = k_*((\delta_i)_{i=0}^{N-1}, (y_i^\delta)_{i=0}^{N-1})$ such that

$$k_* := \min\{lN : l \in \mathbb{N} \text{ and } x_{lN}^\delta = x_{lN+1}^\delta = \cdots = x_{lN+N}^\delta\}. \tag{1.8}$$

In other words, $k_* \in \mathbb{N}$ is the smallest multiple of N such that $x_{k_*}^\delta = x_{k_*+1}^\delta = \cdots = x_{k_*+N}^\delta$.

Outline of the manuscript

The article is organized as follows: In Section 2 we introduce the rrLMK method, proposed and analyzed in this manuscript. A detailed formulation of this method is given (Section 2.2). Moreover, we prove that the method is well defined (Section 2.3). In Section 3 a convergence analysis for the rrLMK method is presented. The main results discussed in this section are: convergence in the exact data case (Theorem 3.9), stability and semi-convergence results (Theorems 3.11 and 3.12, respectively). Section 4 is devoted to numerical experiments. Here we consider the Complete Electrode Model for Electrical Impedance Tomography (EIT-CEM) [24]. For the reconstructions we use both synthetic data (Section 4.2) and real data (Section 4.3). The performance of the rrLMK method is compared with other Kaczmarz-type methods, namely the geometric LMK method (gLMK) with α_k chosen in geometric progression, the stationary LMK method (sLMK) with constant α_k. Section 5 is dedicated to final remarks and conclusions.

2 A range-relaxed Levenberg–Marquardt–Kaczmarz method

In the sequel we introduce the *range-relaxed Levenberg–Marquardt–Kaczmarz* (rrLMK) method for solving the ill-posed linear system (1.1)–(1.2). Section 2.1 is devoted to main assumptions needed in the analysis. The new method is presented in Section 2.2 and a corresponding algorithm is discussed. In Section 2.3 we prove that, under the main assumptions, the rrLMK method is well defined.

The implementable method proposed here happens to be a LMK-type method where, in each iteration, the set of feasible choices for the Lagrange multipliers is an interval, instead of a single real number. For this reason, this method is called range-relaxed Levenberg–Marquardt–Kaczmarz method.

2.1 Main assumptions

Throughout this article we assume that the intersection of the domains of definition $D(F_i)$ have non-empty interior, and that the initial guess $x_0 \in X$ satisfies $B_\rho(x_0) \subset \bigcap_{i=0}^{N-1} D(F_i)$ for some $\rho > 0$. Additionally:

(A1) Each operator F_i is Fréchet differentiable with continuous derivative F_i'. Moreover, there exists a constant $C > 0$ such that

$$\|F_i'(x)\| \leq C \quad \text{for } i = 0, \dots, N-1 \text{ and for all } x \in B_\rho(x_0). \tag{2.1}$$

(A2) The *weak Tangential Cone Condition* (wTCC) holds at $B_\rho(x_0)$, with $0 < \eta < 1$, i.e.,

$$\|F_i(\bar{x}) - F_i(x) - F_i'(x)(\bar{x} - x)\|_Y \leq \eta \|F_i(\bar{x}) - F_i(x)\|_Y \quad \text{for } i = 0, \dots, N-1 \text{ and for all } x, \bar{x} \in B_\rho(x_0). \tag{2.2}$$

(A3) There exists $x^\star \in B_{\rho/2}(x_0)$ such that $F_i(x^\star) = y_i$, $i = 0, \dots, N-1$, where $y_i \in \mathrm{Rg}(F_i)$ are the exact data satisfying (1.1), i.e., $x^\star$ is a (non-necessarily unique) solution of (1.2).

2.2 Description of the method

As mentioned in the introduction, the iterative step of the rrLMK method is inspired in the step proposed in [18]. In what follows we address the computation of the iterate x_{k+1}^δ, given x_k^δ. From this point on we adopt the simplified notation $i := [k]$.

If the iterate x_k^δ is such that $\|F_i(x_k^\delta) - y_i^\delta\| \leq \tau\delta_i$, then we set $x_{k+1}^\delta := x_k^\delta$. Otherwise, we define for each $\mu > 0$ the set

$$\Omega_\mu^i := \{x \in X : \|y_i - F_i(x_k^\delta) - F_i'(x_k^\delta)(x - x_k^\delta)\| \leq \mu\}$$

(the careful reader observes that Ω_μ^i, $\mu > 0$, are levelsets of the residual (linearized at $x = x_k^\delta$) of the ith equation of the nonlinear system (1.2)). The next iterate x_{k+1}^δ is obtained by solving for $(x, \mu) \in X \times \mathbb{R}$ the *range-relaxed projection problem*

$$\begin{cases} \min\limits_x \|x - x_k^\delta\|^2 \\ \text{s.t. } \|y_i - F_i(x_k^\delta) - F_i'(x_k^\delta)(x - x_k^\delta)\|^2 \leq \mu^2 \\ \text{and } \bar{\Phi}(\|F_i(x_k^\delta) - y_i^\delta\|, \delta_i) \leq \mu \leq \bar{\bar{\Phi}}(\|F_i(x_k^\delta) - y_i^\delta\|, \delta_i), \end{cases} \tag{2.3}$$

where the functions $\bar{\Phi}, \bar{\bar{\Phi}} : \mathbb{R}^2 \to \mathbb{R}$ are defined by

$$\bar{\Phi}(u, v) = (\bar{p} + \eta)u + (\eta + 1 - \bar{p})v \quad \text{and} \quad \bar{\bar{\Phi}}(u, v) = (\bar{\bar{p}} + \eta)u + (\eta + 1 - \bar{\bar{p}})v \quad \text{for all } u, v \in \mathbb{R},$$

with $0 < \bar{p} < \bar{\bar{p}} < 1$ chosen appropriately (see Algorithm 1). Notice that the interval

$$[\bar{\Phi}(\|F_i(x_k^\delta) - y_i^\delta\|, \delta_i), \bar{\bar{\Phi}}(\|F_i(x_k^\delta) - y_i^\delta\|, \delta_i)] \subset \mathbb{R}^+$$

is non-degenerate whenever $\|F_i(x_k^\delta) - y_i^\delta\| > \delta_i$.

Once a solution (x', μ') of (2.3) has been computed, define $x_{k+1}^\delta := x'$ (consequently, the linearized residual satisfies $\|y_i^\delta - F_i(x_k^\delta) - F_i'(x_k^\delta)(x_{k+1}^\delta - x_k^\delta)\| = \mu'$). It is worth mentioning that x_{k+1}^δ is generated from x_k^δ by projecting it onto any one of the family of closed convex sets $(\Omega_\mu^i)_{\bar{\Phi} \leq \mu \leq \bar{\bar{\Phi}}}$ (see, e.g., [18] or [6, Section 5.1]).

In Algorithm 1 we present the *rrLMK method* in algorithmic form. The above discussed range-relaxed projection problem can be recognized in Step 3.2.

Remark 2.1. Some remarks on Algorithm 1:

(a) In Step 3.2 there is a safeguard to guarantee $\alpha_k \geq \alpha_{\min}$. This fact is used in the proof of convergence of the rrLMK (Theorem 3.9).

(b) In the (realistic) noisy data case, a lower bound on α_k is not required for the implementation of Algorithm 1. Indeed, since $\delta_i > 0$ is fixed and the stopping criteria is always reached after a finite number of steps (Proposition 3.6), the constant $\alpha_{\min} > 0$ can be chosen sufficiently small such that $\alpha_k \geq \alpha_{\min}$ for $k = 0, \dots, k_*$.

Algorithm 1. The range-relaxed Levenberg–Marquardt–Kaczmarz (rrLMK) method

(1) Choose an initial guess $x_0 \in X$ and $\alpha_{\min} > 0$; set $k := 0$.

(2) Choose the positive constants τ, $\bar{p}$ and $\bar{\bar{p}}$ s.t.

$$\tau > (1+\eta)(1-\eta)^{-1} \quad \text{and} \quad 0 < \bar{p} < \bar{\bar{p}} < [\tau(1-\eta) - (1+\eta)](\tau - 1)^{-1}. \tag{2.4}$$

(3) **repeat**

(3.1) $i = [k]$;

(3.2) **if** $[\|F_i(x_k^\delta) - y_i^\delta\| > \tau\delta_i]$ **then**

compute $(\alpha_k, h_k^\delta) \in \mathbb{R}^+ \times X$ such that

$$h_k^\delta = (F_i'(x_k^\delta)^* F_i'(x_k^\delta) + \alpha_k I)^{-1} F_i'(x_k^\delta)^* (y_i^\delta - F_i(x_k^\delta)), \tag{2.5}$$

$$\|y_i^\delta - F_i(x_k^\delta) - F_i'(x_k^\delta) h_k^\delta\| \in [c_k, d_k], \tag{2.6}$$

where

$$c_k = (\bar{p} + \eta)\|F_i(x_k^\delta) - y_i^\delta\| + (\eta + 1 - \bar{p})\delta_i, \tag{2.7}$$

$$d_k = (\bar{\bar{p}} + \eta)\|F_i(x_k^\delta) - y_i^\delta\| + (\eta + 1 - \bar{\bar{p}})\delta_i. \tag{2.8}$$

if $[\alpha_k < \alpha_{\min}]$ **then**

$\alpha_k = \alpha_{\min}$; $h_k^\delta = (F_i'(x_k^\delta)^* F_i'(x_k^\delta) + \alpha_{\min} I)^{-1} F_i'(x_k^\delta)^* (y_i^\delta - F_i(x_k^\delta))$;

endif

else

$h_k^\delta = 0$;

endif

(3.3) $x_{k+1}^\delta = x_k^\delta + h_k^\delta$; $k = k + 1$;

until $[([k] = 0)$ and $(h_{k-1}^\delta = h_{k-2}^\delta = \cdots = h_{k-N}^\delta = 0)]$;

(4) $k_* = k - N$.

(c) The inequality $\|y_i^\delta - F_i(x_k^\delta) - F_i'(x_k^\delta)h_k^\delta\| \geq c_k$ holds in both cases $\alpha_k > \alpha_{\min}$ and $\alpha_k = \alpha_{\min}$ (in the last case this inequality follows from Algorithm 1 and Lemma 2.2 (2) and (4)). However, the inequality $\|y_i^\delta - F_i(x_k^\delta) - F_i'(x_k^\delta)h_k^\delta\| \leq d_k$ can only be guaranteed in the case $\alpha_k > \alpha_{\min}$.

(d) The Lagrange multipliers α_k are defined if and only if $\|F_i(x_k^\delta) - y_i^\delta\| > \tau\delta_i$ holds. Otherwise, we apply the loping strategy proposed in [10] and set $x_{k+1} := x_k$, avoiding the computation of the pair (α_k, h_k^δ).

(e) The inequalities $(1+\eta)(1-\eta)^{-1} > 1$, $[\tau(1-\eta) - (1+\eta)](\tau-1)^{-1} < 1$, $\bar{p} + \eta < 1$ (see (2.4)) are used several times in the manuscript.

2.3 Preliminary results

The remaining of this section is devoted to verify that, under assumptions (A1), (A2) and (A3), Step 3.2 in Algorithm 1 is well defined (see Theorem 2.4). We start the discussion with a lemma, which contains a collection of results related to the solvability of the *range-relaxed projection problem* (2.3). For a proof, we refer the reader to [18, Lemma 2.3].

Lemma 2.2. *Let $A : X \to Y$ be a continuous linear mapping, $\bar{z} \in X$ and $b \in Y$ a vector with non-zero projection onto the range of A. Define, for $\alpha > 0$,*

$$z_\alpha := \arg\min_{z \in X} \|A(z - \bar{z}) - b\|^2 + \alpha\|z - \bar{z}\|^2. \tag{2.9}$$

The following assertions hold:

(1) $z_\alpha = \bar{z} + (A^*A + \alpha I)^{-1} A^* b$,

(2) $\alpha \mapsto \|A(z_\alpha - \bar{z}) - b\|$ *is a continuous, strictly increasing function on* $\alpha > 0$,

(3) $\lim_{\alpha\to 0}\|A(z_\alpha - \bar{z}) - b\| = \inf_{z\in X}\|A(z-\bar{z}) - b\|$,
(4) $\lim_{\alpha\to\infty}\|A(z_\alpha - \bar{z}) - b\| = \|b\|$,
(5) $\alpha \le \|A^* b\|^2[\|b\|(\|b\| - \|A(z_\alpha - \bar{z}) - b\|)]^{-1}$.

The next result will allow us to compare the non-linear residual $\|y_i^\delta - F_i(x)\|$ with the linearized residual $\|y_i^\delta - F_i(x) - F_i'(x)(x^* - x)\|$ for $x \in B_\rho(x_0)$ and $x^* \in B_\rho(x_0)$ a solution of (1.2).

Lemma 2.3. *Let assumptions* (A1)–(A2) *hold, and let* $x^* \in B_\rho(x_0)$ *be a solution of problem* (1.2). *Then, for* $i = 0, \dots, N-1$, *we have*

$$\|y_i^\delta - F_i(x) - F_i'(x)(x^* - x)\| \le \eta\|y_i^\delta - F_i(x)\| + (1+\eta)\delta_i \quad \text{for all } x \in B_\rho(x_0).$$

Moreover, in the exact data case, $\|y_i - F_i(x) - F_i'(x)(x^* - x)\| \le \eta\|y_i - F_i(x)\|$ *for all* $x \in B_\rho(x_0)$.

Proof. The proof follows the lines of the proof of [18, Lemma 2.4]. □

Notice that, if assumptions (A1)–(A3) hold, then Lemma 2.3 is valid for $x^* = x^\star$ as in (A3). The next result guarantees the well definedness of Step 3.2 in Algorithm 1.

Theorem 2.4. *Let assumptions* (A1)–(A3) *hold, and let* (x_k^δ) *be a sequence generated by Algorithm 1. If, for some* $k \ge 0$, $x_k^\delta \in B_\rho(x_0)$ *and* $\|F_{[k]}(x_k^\delta) - y_{[k]}^\delta\| > \tau\delta_{[k]}$, *then there exists a pair* $(\alpha_k \in \mathbb{R}^+, h_k^\delta \in X)$ *solving* (2.5)–(2.6).

Proof. We first consider the noisy data case. It follows from the assumption $\|F_i(x_k^\delta) - y_i^\delta\| > \tau\delta_i$ that $c_k < d_k$. Thus, we can define the set $J := \{\alpha > 0 : \|y_i^\delta - F_i(x_k^\delta) - F_i'(x_k^\delta)(\xi_\alpha - x_k^\delta)\| \in [c_k, d_k]\}$, where ξ_α is defined by the mapping

$$(0, +\infty) \ni \alpha \mapsto \xi_\alpha := \arg\min_{\xi\in X}\|y_i^\delta - F_i(x_k^\delta) - F_i'(x_k^\delta)(\xi - x_k^\delta)\|^2 + \alpha\|\xi - x_k^\delta\|^2 \in X.$$

To complete the proof in the noisy data case, it is enough to verify that J is a non-empty, non-degenerate interval.

Notice that $c_k = \bar{p}(\|F_i(x_k^\delta) - y_i^\delta\| - \delta_i) + \eta\|F_i(x_k^\delta) - y_i^\delta\| + (1+\eta)\delta_i > \eta\|F_i(x_k^\delta) - y_i^\delta\| + (1+\eta)\delta_i$. Therefore, it follows from Lemma 2.3 (with $x = x_k^\delta$ and $x^* = x^\star$ as in (A3)) that

$$c_k > \|y_i^\delta - F_i(x_k^\delta) - F_i'(x_k^\delta)(x^\star - x_k^\delta)\| \ge \inf_{z\in X}\|y_i^\delta - F_i(x_k^\delta) - F_i'(x_k^\delta)(z - x)\|.$$

On the other hand, it follows from (2.4) that $[\bar{\bar{p}} + \eta + \frac{\eta+1-\bar{\bar{p}}}{\tau}] < 1$. Consequently,

$$d_k = (\bar{\bar{p}} + \eta)\|F_i(x_k^\delta) - y_i^\delta\| + (\eta + 1 - \bar{\bar{p}})\delta_i < \left[\bar{\bar{p}} + \eta + \frac{\eta + 1 - \bar{\bar{p}}}{\tau}\right]\|F_i(x_k^\delta) - y_i^\delta\| < \|F_i(x_k^\delta) - y_i^\delta\|.$$

Summarizing our findings:

$$\inf_z\|y_i^\delta - F_i(x_k^\delta) - F_i'(x_k^\delta)(z - x)\| < c_k < d_k < \|F_i(x_k^\delta) - y_i^\delta\|.$$

The existence of (α_k, h_k^δ) solving (2.5), (2.6) follows from these inequalities and Lemma 2.2 (2)–(4).

Next we consider the exact data case. The proof is analogous to the noisy data case. In order to estimate d_k, one uses the inequality $\bar{\bar{p}} + \eta < 1$ (see (2.4)) to conclude that $d_k < \|F_i(x_k) - y_i\|$. □

3 Convergence analysis

We start this section deriving estimates for the "gain" $\|x^* - x_k^\delta\|^2 - \|x^* - x_{k+1}^\delta\|^2$, where (x_k^δ) is a sequence generated by the rrLMK method.

Proposition 3.1. *Let assumptions* (A1)–(A2) *hold, and let* (x_k^δ) *be a sequence generated by Algorithm 1 in the noisy data case. Moreover, let* $x^* \in B_\rho(x_0)$ *be a solution of* (1.2). *If* $x_k^\delta \in B_\rho(x_0)$ *for some* $k \ge 0$, *two scenarios may occur:*

(i) *If* $\|F_i(x_k^\delta) - y_i^\delta\| > \tau\delta_i$, *then*

$$\begin{aligned}&\|x^* - x_k^\delta\|^2 - \|x^* - x_{k+1}^\delta\|^2\\ &\quad\geq \|x_{k+1}^\delta - x_k^\delta\|^2 + 2\bar{p}\Big(1 - \frac{1}{\tau}\Big)\alpha_k^{-1}\|y_i^\delta - F_i(x_k^\delta) - F_i'(x_k^\delta)(x_{k+1}^\delta - x_k^\delta)\|\|y_i^\delta - F_i(x_k^\delta)\|.\end{aligned} \tag{3.1}$$

(ii) *If* $\|F_i(x_k^\delta) - y_i^\delta\| \leq \tau\delta_i$, *then* $\|x^* - x_k^\delta\|^2 - \|x^* - x_{k+1}^\delta\|^2 = 0$.
Additionally, in the exact data case, if $\|F_i(x_k) - y_i\| = 0$ *we have* $\|x^* - x_{k+1}\| = \|x^* - x_k\|$, *otherwise*

$$\|x^* - x_k\|^2 - \|x^* - x_{k+1}\|^2 \geq \|x_{k+1} - x_k\|^2 + 2\bar{p}\alpha_k^{-1}\|y_i - F_i(x_k) - F_i'(x_k)(x_{k+1} - x_k)\|\|y_i - F_i(x_k)\|. \tag{3.2}$$

Proof. Item (ii) follows immediately from Algorithm 1, since $\|F_i(x_k^\delta) - y_i^\delta\| \leq \tau\delta_i$ implies $x_{k+1}^\delta = x_k^\delta$. In order to prove Item i), we derive from the polarization formula the identity

$$\|x^* - x_k^\delta\|^2 - \|x^* - x_{k+1}^\delta\|^2 = \|x_{k+1}^\delta - x_k^\delta\|^2 - 2\langle x_{k+1}^\delta - x_k^\delta, x_{k+1}^\delta - x^*\rangle. \tag{3.3}$$

We adopt the notation $A_k = F_i'(x_k^\delta)$, $b_k^\delta = y_i^\delta - F_i(x_k^\delta)$. From $h_k^\delta = \alpha_k^{-1}A_k^*(b_k^\delta - A_k h_k^\delta)$ follows

$$\begin{aligned}-\langle x_{k+1}^\delta - x_k^\delta, x_{k+1}^\delta - x^*\rangle &= \alpha_k^{-1}\langle A_k^*(A_k h_k^\delta - b_k^\delta), x_{k+1}^\delta - x^*\rangle\\ &= \alpha_k^{-1}\langle A_k h_k^\delta - b_k^\delta, A_k(h_k^\delta - (x^* - x_k^\delta)) + b_k^\delta - b_k^\delta\rangle\\ &= \alpha_k^{-1}\big[\langle A_k h_k^\delta - b_k^\delta, A_k h_k^\delta - b_k^\delta\rangle - \langle A_k h_k^\delta - b_k^\delta, A_k(x^* - x_k^\delta) - b_k^\delta\rangle\big]\\ &\geq \alpha_k^{-1}\|A_k h_k^\delta - b_k^\delta\|\big[\|A_k h_k^\delta - b_k^\delta\| - \|A_k(x^* - x_k^\delta) - b_k^\delta\|\big].\end{aligned} \tag{3.4}$$

In order to estimate the last term in (3.4), we derive from (2.7) the identity $(1+\eta)\delta_i + \eta\|b_k^\delta\| = c_k + \bar{p}\delta_i - \bar{p}\|b_k^\delta\|$. Thus, we conclude from Lemma 2.3, Algorithm 1 and $\|b_k^\delta\| > \tau\delta_i$ that

$$\begin{aligned}\|A_k(x^* - x_k^\delta) - b_k^\delta\| &\leq (1+\eta)\delta_i + \eta\|b_k^\delta\| = c_k + \bar{p}\delta_i - \bar{p}\|b_k^\delta\|\\ &< \|A_k h_k^\delta - b_k^\delta\| + \bar{p}\Big(\frac{1}{\tau} - 1\Big)\|b_k^\delta\|\end{aligned} \tag{3.5}$$

(in the last inequality we used the fact $c_k \leq \|A_k h_k^\delta - b_k^\delta\|$, see Remark 2.1 (c)). Consequently, (3.1) follows by substituting (3.5) and (3.4) in (3.3).

What concerns the exact data case, the first assertion is trivial. On the other hand, if $\|F_i(x_k) - y_i\| > 0$, we argue as in the noisy data case in order to obtain, instead of (3.5), the inequalities

$$\|A_k(x^* - x_k) - b_k\| \leq \eta\|b_k\| = c_k - \bar{p}\|b_k\| \leq \|A_k h_k - b_k\| - \bar{p}\|b_k\| \tag{3.6}$$

(here $b_k := y_i - F_i(x_k)$). Then the derivation of (3.2) is analogous to the proof of inequality (3.1) in the noisy data case. □

The next result is a direct consequence of Proposition 3.1. It guarantees that any sequence generated by Algorithm 1 does not leave the ball $B_\rho(x_0)$.

Corollary 3.2. *Let assumptions* (A1)–(A3) *hold, and let* (x_k^δ) *be a sequence generated by Algorithm 1 in the noisy data case. Then* $x_k^\delta \in B_\rho(x_0)$ *for* $k = 0, \ldots, k_* - 1$. *Moreover,*

$$\|x^\star - x_{k+1}^\delta\| \leq \|x^\star - x_k^\delta\|, \quad k = 0, \ldots, k_* - 1, \tag{3.7}$$

where $x^\star$ *is given as in* (A3). *Additionally, in the exact data case we have* $(x_k) \in B_\rho(x_0)$, *and*

$$\|x^\star - x_{k+1}\| \leq \|x^\star - x_k\| \quad \text{for } k = 0, 1, \ldots .$$

The careful reader observes that, for any $x^* \in B_\rho(x_0)$ solution of (1.2) it holds $\|x^* - x_{k+1}^\delta\| \leq \|x^* - x_k^\delta\|$, $k = 0, \ldots, k_* - 1$. Moreover, in the exact data case $\|x^* - x_{k+1}\| \leq \|x^* - x_k\|$, for $k = 0, 1, \ldots$

Lemma 3.3. *Let assumptions* (A1), (A2), (A3) *hold, and let* (x_k^δ) *be a sequence generated by Algorithm 1. If* $\|F_i(x_k^\delta) - y_i^\delta\| > \tau\delta_i$, *then*

$$\alpha_k \leq \alpha_{\max} := \max\Big\{\frac{C^2}{(1-\bar{p})(1-\tau^{-1}) - \eta(1+\tau^{-1})}, \alpha_{\min}\Big\}. \tag{3.8}$$

Additionally, in the exact data case, if $\|F_i(x_k) - y_i\| > 0$, *then* $\alpha_k \leq \alpha_{\max} := \max\{\frac{C^2}{1-(\bar{p}+\eta)}, \alpha_{\min}\}$.

Proof. Assume $\alpha_k > \alpha_{\min}$. From Lemma 2.2 (5) with $\alpha = \alpha_k$, $\overline{z} = x_k^\delta$, $z_\alpha = x_{k+1}^\delta$, $b = y_i^\delta - F_i(x_k^\delta)$ and $A = F_i'(x_k^\delta)$ it follows that

$$\begin{aligned}\alpha_k &\le \frac{\|F_i'(x_k^\delta)^*(y_i^\delta - F_i(x_k^\delta))\|^2}{\|y_i^\delta - F_i(x_k^\delta)\|(\|y_i^\delta - F_i(x_k^\delta)\| - \|y_i^\delta - F_i(x_k^\delta) - F_i'(x_k^\delta)h_k^\delta\|)} \\ &\le \frac{C^2\|y_i^\delta - F_i(x_k^\delta)\|}{\|y_i^\delta - F_i(x_k^\delta)\| - \|y_i^\delta - F_i(x_k^\delta) - F_i'(x_k^\delta)h_k^\delta\|}. \end{aligned} \tag{3.9}$$

Moreover, it follows from Algorithm 1 (see (2.6) and (2.8)) that for all $k = 0, 1, \dots, k_* - 1$,

$$\begin{aligned}\|y_i^\delta - F_i(x_k^\delta)\| - \|y_i^\delta - F_i(x_k^\delta) - F_i'(x_k^\delta)h_k^\delta\| &\ge \|y_i^\delta - F_i(x_k^\delta)\| - d_k \\ &\ge \Big(1 - (\bar{\bar{p}} + \eta) - \frac{\eta + 1 - \bar{\bar{p}}}{\tau}\Big)\|y_i^\delta - F_i(x_k^\delta)\| \\ &= \big((1 - \bar{\bar{p}})(1 - \tau^{-1}) - \eta(1 + \tau^{-1})\big)\|y_i^\delta - F_i(x_k^\delta)\| \end{aligned} \tag{3.10}$$

(notice that $(1 - \bar{\bar{p}})(1 - \tau^{-1}) - \eta(1 + \tau^{-1}) > 0$, see (2.4)). Thus, (3.8) follows from (3.9) and (3.10).

Next we address the exact data case. In this case $d_k = (\bar{\bar{p}} + \eta)\|y_i - F_i(x_k)\|$. Once again we assume $\alpha_k > \alpha_{\min}$. Consequently,

$$\|y_i - F_i(x_k)\| - \|y_i - F_i(x_k) - F_i'(x_k)h_k\| \ge (1 - (\bar{\bar{p}} + \eta))\|y_i - F_i(x_k)\|$$

(notice that $1 - (\bar{\bar{p}} + \eta) > 0$, see (2.4)). From this inequality and (3.9) follows $\alpha_k \le C^2[1 - (\bar{\bar{p}} + \eta)]^{-1}$. Therefore, it holds $\alpha_k \le \max\{\frac{C^2}{1-(\bar{\bar{p}}+\eta)}, \alpha_{\min}\}$ completing the proof. □

Lemma 3.4. *Let assumptions* (A1)–(A3) *hold, and let* (x_k^δ) *be a sequence generated by Algorithm 1. For* $k = 0, 1, \dots, k_* - 1$ *we have*

$$(\bar{\bar{p}} + \eta)\|y_i^\delta - F_i(x_k^\delta)\| \le \|y_i^\delta - F_i(x_k^\delta) - F_i'(x_k^\delta)(x_{k+1}^\delta - x_k^\delta)\| \le \|y_i^\delta - F_i(x_k^\delta)\|. \tag{3.11}$$

In the exact data case (3.11) *holds for* $k = 0, 1, \dots$.

Proof. First we address the exact data case. We adopt the notation

$$A_k = F_i'(x_k), \quad b_k = y_i - F_i(x_k) \quad \text{and} \quad h_k = x_{k+1} - x_k.$$

From Corollary 3.2 it follows that $x_k \in B_\rho(x_0)$ for $k \ge 0$. If $\|F_i(x_k) - y_i\| = 0$, then $x_{k+1} = x_k$ and (3.11) is trivial. On the other hand, if $\|F_i(x_k) - y_i\| > 0$, the second inequality in (3.11) follows from Lemma 2.2 (2) and (4) (with $\alpha = \alpha_k$, $\overline{z} = x_k$, $z_\alpha = x_{k+1}$, $b = b_k$ and $A = A_k$). Moreover, it follows from Remark 2.1 (c) and the definition of c_k in (2.7)

$$(\eta + \bar{p})\|b_k\| = c_k \le \|A_k h_k - b_k\|, \tag{3.12}$$

which is the first inequality in (3.11).

In the noisy data case, the proof is analogous. The main difference is that, if $\delta_i > 0$, (3.12) is replaced by

$$(\eta + \bar{p})\|b_k^\delta\| \le c_k \le \|A_k h_k^\delta - b_k^\delta\|$$

(where $b_k^\delta = y_i^\delta - F_i(x_k^\delta)$ and $h_k^\delta = x_{k+1}^\delta - x_k^\delta$). □

Proposition 3.5. *Let assumptions* (A1), (A2) *and* (A3) *hold. Any sequence* (x_k) *generated by Algorithm 1 in the exact data case satisfies*

$$\sum_{k=0}^{\infty} \alpha_k^{-1}\|y_i - F_i(x_k) - F_i'(x_k)h_k\|\|y_i - F_i(x_k)\| < \infty, \qquad \sum_{k=0}^{\infty} \|x_{k+1} - x_k\|^2 < \infty,$$

$$\sum_{k=0}^{\infty} \|y_i - F_i(x_k)\|^2 < \infty, \quad \sum_{k=0}^{\infty} \|y_i - F_i(x_k) - F_i'(x_k)h_k\|^2 < \infty.$$

Proof. The summability of the first two series follow from Proposition 3.1. The summability of the third and fourth series follows from the summability of the first one, together with Lemma 3.3 and Lemma 3.4. □

Proposition 3.6. *Let assumptions* (A1)–(A3) *hold, and let* (x_k^δ) *be a sequence generated by Algorithm 1. The stopping index* k_* *defined in* (1.8) *is finite and satisfies*

$$k_* \leq N\Big(1 + \frac{\alpha_{\max}\|x^\star - x_0\|^2}{2\bar{p}(\tau-1)(\bar{p}+\eta)\tau\delta_{\min}^2}\Big) \tag{3.13}$$

with $\alpha_{\max} > 0$ *defined as in Lemma 3.3.*

Proof. Assume by contradiction that k^* is not finite. Thus, in each cycle $l \in \mathbb{N}$ of the rrLMK method (consisting of steps $\{lN, lN+1, \dots, lN+N-1\}$) there exists at least one index $j(l) \in \{0, \dots, N-1\}$ such that the inequality $\|F_{j(l)}(x^\delta_{lN+j(l)}) - y^\delta_{j(l)}\| > \tau\delta_{j(l)}$ holds. Consequently, it follows from Proposition 3.1 that (3.1) holds for $k = lN + j(l)$, with $l \in \mathbb{N}$. Moreover, we know from Proposition 3.1 and Algorithm 1 that either (3.1) holds or $x^\delta_{k+1} = x^\delta_k$. Therefore, for all $l \in \mathbb{N}$ we have

$$\begin{aligned}
\|x^\star - x_0\|^2 &\geq \sum_{k=0}^{lN} \|x^\star - x_k^\delta\|^2 - \|x^\star - x_{k+1}^\delta\|^2 \geq \sum_{s=0}^{l} \|x^\star - x^\delta_{sN+j(s)}\|^2 - \|x^\star - x^\delta_{sN+j(s)+1}\|^2 \\
&\geq 2\bar{p}\Big(1 - \frac{1}{\tau}\Big) \sum_{s=0}^{l} \alpha^{-1}_{sN+j(s)} \|y^\delta_{j(s)} - F_{j(s)}(x^\delta_{sN+j(s)}) - F'_{j(s)}(x^\delta_{sN+j(s)})(x^\delta_{sN+j(s)+1} - x^\delta_{sN+j(s)})\| \\
&\qquad \cdot \|y^\delta_{j(s)} - F_{j(s)}(x^\delta_{sN+j(s)})\| \\
&\geq 2\bar{p}\Big(1 - \frac{1}{\tau}\Big)\alpha^{-1}_{\max} \sum_{s=0}^{l} (\bar{p}+\eta)\|y^\delta_{j(s)} - F_{j(s)}(x^\delta_{sN+j(s)})\|^2 \\
&\geq 2l\bar{p}\Big(1 - \frac{1}{\tau}\Big)\alpha^{-1}_{\max}(\bar{p}+\eta)\tau^2\delta^2_{\min}
\end{aligned} \tag{3.14}$$

(to derive the fourth inequality we used Lemma 3.3 and Lemma 3.4). Since the right-hand side of (3.14) becomes unbounded as $l \to \infty$, a contradiction is established and the finiteness of k_* follows. To complete the proof, we derive (3.13) by substituting $l = \frac{k^*-N}{N}$ in (3.14). □

In what follows we discuss the evolution of the non-linear residual in the exact data case.

Remark 3.7. Let assumptions (A1)–(A3) hold, and let (x_k) be a sequence generated by Algorithm 1 in the exact data case. If, for some $k \geq 0$, $\|F_i(x_k) - y_i\| > 0$, then it holds

$$\|y_i - F_i(x_{k+1})\| \leq \Lambda_1 \|y_i - F_i(x_k)\|, \tag{3.15}$$

where $\Lambda_1 = \sqrt{2 + 2C^4\alpha^{-2}_{\min}(1-\eta)^{-2}}$.

Proof. It follows from (A2) that

$$\begin{aligned}
\|y_i - F_i(x_{k+1})\|^2 &\leq 2\|y_i - F_i(x_k)\|^2 + 2\|F_i(x_{k+1}) - F_i(x_k)\|^2 \\
&\leq 2\|y_i - F_i(x_k)\|^2 + \frac{2}{(1-\eta)^2}\|F'_i(x_k)(x_{k+1} - x_k)\|^2 \\
&\leq 2\|y_i - F_i(x_k)\|^2 + \frac{2C^2}{(1-\eta)^2}\|x_{k+1} - x_k\|^2.
\end{aligned} \tag{3.16}$$

Moreover, it follows from Step 3.2 of Algorithm 1 that

$$\|x_{k+1} - x_k\|^2 = \|(F'_i(x_k)^* F'_i(x_k) + \alpha_k I)^{-1} F'_i(x_k)^*(y_i - F_i(x_k))\|^2 \leq C^2\alpha^{-2}_{\min}\|y_i - F_i(x_k)\|^2. \tag{3.17}$$

Inequality (3.15) follows now from (3.16) and (3.17). □

We are now ready to state an prove the first main result of this section, namely convergence of the rrLMK method for exact data (Theorem 3.9). First however, we briefly recall the concept of minimum norm solutions.

Definition 3.8. An element $x^\dagger \in X$ is called a x_0-minimal norm solution of problem (1.2) whenever $\|x^\dagger - x_0\| = \inf\{\|x^* - x_0\|, x^*\}$ is a solution of (1.2)$\}$.

Existence and uniqueness of x_0-minimal norm solutions of (1.2) is guaranteed by [6, 15]. Let assumptions (A1), (A2) and (A3) hold. Given $x^* \in B_{\rho/2}(x_0)$ a solution of (1.2) and $z \in N(F_i'(x^*))$, for some $0 \le i \le N-1$, then the element $x^* + tz \in B_\rho(x_0)$ is also a solution of $F_i(x) = y_i$ for all $t \in (-\frac{\rho}{2}, \frac{\rho}{2})$. Due to (A3), $x^\dagger \in B_{\rho/2}(x_0)$. Consequently, the inequality

$$\|x^\dagger - x_0\| \le \|(x^\dagger + tz) - x_0\|$$

holds for all $t \in (-\frac{\rho}{2}, \frac{\rho}{2})$ and all $z \in \bigcap_{i=0}^{N-1} N(F_i'(x^*))$, from where we conclude that

$$x^\dagger - x_0 \in \left[\bigcap_{i=0}^{N-1} N(F_i'(x^\dagger))\right]^\perp. \tag{3.18}$$

Theorem 3.9 (Convergence for exact data). *Let assumptions* (A1)–(A3) *hold, and let* (x_k) *be a sequence generated by Algorithm 1 in the exact data case. Either* (x_k) *terminates after finitely many iterations with a solution of* (1.2)*, or* (x_k) *converges to a solution of* (1.2) *as* $k \to \infty$*. Additionally, if*

$$N(F_j'(x^\dagger)) \subset N(F_j'(x)) \quad \text{for all } x \in B_\rho(x_0), \quad j = 0, 1, \dots, N-1, \tag{3.19}$$

holds, then $x_k \to x^\dagger$ *as* $k \to \infty$.

Proof. We use the notation $A_k = F_i(x_k)$, $b_k = y_i - F_i(x_k)$ and $e_k = x^\star - x_k$, for $k \ge 0$.

If the rrLMK method stops after $k_* < \infty$ steps, it follows from Algorithm 1 that x_{k_*} is a solution of (1.2).

Otherwise, it follows from Lemma 3.2 that $(\|e_k\|)$ is monotone non-increasing. Thus, $\|e_k\| \to \varepsilon$, for some $\varepsilon \ge 0$, as $k \to \infty$. In what follows we verify that (e_k) is a Cauchy sequence.

In order to prove that (e_k) is indeed a Cauchy sequence, it suffices to prove that

$$|\langle e_n - e_k, e_n\rangle| \to 0$$

and

$$|\langle e_l - e_n, e_n\rangle| \to 0$$

as $k, l \to \infty$, with $k \le l$ for some $k \le n \le l$ (see, e.g., [12, Theorem 2.3]).

Given $k \le l$ arbitrary, we write

$$k = k_0 N + k_1 \quad \text{and} \quad l = l_0 N + l_1$$

with $k_1, l_1 \in \{0, \dots, N-1\}$, and choose $n_0 \in \{k_0, \dots, l_0\}$ such that

$$\sum_{s=0}^{N-1} \|F_s(x_{n_0 N+s}) - y_s\| \le \sum_{s=0}^{N-1} \|F_s(x_{j_0 N+s}) - y_s\| \tag{3.20}$$

for all $j_0 \in \{k_0, \dots, l_0\}$. Next we define $n := n_0 N + N - 1$ (if $n_0 = l_0$, we set $n := n_0 N + l_1$; so that $k \le n \le l$), and estimate

$$\begin{aligned}
|\langle e_n - e_k, e_n\rangle| &= \left|\sum_{j=k}^{n-1} \langle (x_{j+1} - x_j), (x^\star - x_n)\rangle\right| \\
&= \left|\sum_{j=k}^{n-1} \alpha_j^{-1} \langle A_j h_j - b_j, A_j(x^\star - x_n)\rangle\right| \\
&\le \sum_{j=k}^{n-1} \alpha_j^{-1} \|A_j h_j - b_j\| \|A_j(x^\star - x_n)\| \\
&= \sum_{j=k}^{n-1} \alpha_j^{-1} \|A_j h_j - b_j\| \left[\|F_{[j]}'(x_j)\,(x_n - x_j)\| + \|F_{[j]}'(x_j)(x_j - x^\star)\|\right] \\
&\le (1+\eta) \sum_{j_0=k_0}^{n_0} \sum_{j_1=0}^{N-1} \alpha_j^{-1} \|A_j h_j - b_j\| \|F_{j_1}(x_n) - y_{j_1}\| \\
&\quad + 2(1+\eta) \sum_{j_0=k_0}^{n_0} \sum_{j_1=0}^{N-1} \alpha_j^{-1} \|A_j h_j - b_j\| \|F_{j_1}(x_j) - y_{j_1}\|
\end{aligned} \tag{3.21}$$

(we use the notation $j = j_0 N + j_1$). The term $\|F_{j_1}(x_n) - y_{j_1}\|$ in (3.21) can be estimated by

$$\begin{aligned}
\|F_{j_1}(x_n) - y_{j_1}\| &= \|F_{j_1}(x_{n_0N+N-1}) - y_{j_1}\| \\
&\le \|F_{j_1}(x_{n_0N+j_1}) - y_{j_1}\| + \sum_{s=j_1}^{N-2} \|F_{j_1}(x_{n_0N+s+1}) - F_{j_1}(x_{n_0N+s})\| \\
&\le \|F_{j_1}(x_{n_0N+j_1}) - y_{j_1}\| + \frac{1}{1-\eta} \sum_{s=j_1}^{N-2} \|F'_{j_1}(x_{n_0N+s})(x_{n_0N+s+1} - x_{n_0N+s})\| \\
&\le \|F_{j_1}(x_{n_0N+j_1}) - y_{j_1}\| + \frac{C}{1-\eta} \sum_{s=j_1}^{N-2} \|x_{n_0N+s+1} - x_{n_0N+s}\| \\
&\le \|F_{j_1}(x_{n_0N+j_1}) - y_{j_1}\| + \frac{C}{1-\eta} \sum_{s=0}^{N-1} C\alpha_{\min}^{-1} \|F_s(x_{n_0N+s}) - y_s\| \\
&\le \left(1 + \frac{C^2\alpha_{\min}^{-1}}{1-\eta}\right) \sum_{s=0}^{N-1} \|F_s(x_{n_0N+s}) - y_s\|
\end{aligned}$$

(the forth inequality follows from (3.17)). From this inequality and (3.20) we get

$$\|F_{j_1}(x_n) - y_{j_1}\| \le \left(1 + \frac{C^2\alpha_{\min}^{-1}}{1-\eta}\right) \sum_{s=0}^{N-1} \|F_s(x_{j_0N+s}) - y_s\|.$$

Inserting this last inequality into (3.21) ,we obtain

$$\begin{aligned}
|\langle e_n - e_k, e_n\rangle| &\le (1+\eta)\left(1 + \frac{C^2\alpha_{\min}^{-1}}{1-\eta}\right) \sum_{j_0=k_0}^{n_0} \left[\sum_{j_1=0}^{N-1} \alpha_j^{-1} \|A_j h_j - b_j\|\right] \left[\sum_{j_1=0}^{N-1} \|F_{j_1}(x_j) - y_{j_1}\|\right] \\
&\quad + 2(1+\eta) \sum_{j_0=k_0}^{n_0} \sum_{j_1=0}^{N-1} \alpha_j^{-1} \|A_j h_j - b_j\| \, \|F_{j_1}(x_j) - y_{j_1}\| \\
&\le (1+\eta)\left(1 + \frac{C^2\alpha_{\min}^{-1}}{1-\eta}\right) \frac{N}{2} \sum_{j_0=k_0}^{n_0} \left[\alpha_{\min}^{-2} \sum_{j_1=0}^{N-1} \|A_j h_j - b_j\|^2 + \sum_{j_1=0}^{N-1} \|F_{j_1}(x_j) - y_{j_1}\|^2\right] \\
&\quad + 2(1+\eta) \sum_{j_0=k_0}^{n_0} \sum_{j_1=0}^{N-1} \alpha_j^{-1} \|A_j h_j - b_j\| \|F_{j_1}(x_j) - y_{j_1}\|
\end{aligned}$$

(in the last inequality we used the fact $\alpha_k \ge \alpha_{\min}$). Hence, from Proposition 3.5, we have $|\langle e_n - e_k, e_n\rangle| \to 0$ as $k, l \to 0$. Analogously, one proves that $|\langle e_l - e_n, e_n\rangle| \to 0$ as $k, l \to 0$. Therefore, (e_k) is a Cauchy sequence and x_k converges to some $x^+ \in X$. Since the residuals $\|F_i(x_k) - y_i\|$ converge to zero as $k \to \infty$ (see Proposition 3.5), we conclude that x^+ is a solution of (1.2).

To prove the last assertion notice that, if (3.19) holds, then for $k = 0, 1, \ldots,$

$$x_{k+1} - x_k = \alpha_k^{-1} A_k^*(b_k - A_k h_k) \in \mathrm{Rg}(F'_{[k]}(x_k)^*) \subset N(F'_{[k]}(x_k))^\perp \subset N(F'_{[k]}(x^\dagger))^\perp \subset \left[\bigcap_{j=0}^{N-1} N(F'_j(x^\dagger))\right]^\perp.$$

Thus, we conclude

$$x_k - x_0 \in \left[\bigcap_{j=0}^{N-1} N(F'_j(x^\dagger))\right]^\perp.$$

From (3.18), it follows $x_k - x^\dagger \in [\bigcap_{j=0}^{N-1} N(F'_j(x^\dagger))]^\perp$. Consequently,

$$x^+ - x^\dagger = \lim_k (x_k - x^\dagger) \in \left[\bigcap_{j=0}^{N-1} N(F'_j(x^\dagger))\right]^\perp,$$

from where we obtain $x^+ = x^\dagger$. □

Next we derive the second main result of this section, namely the stability property of the rrLMK method. First however, we introduce two relevant concepts.

Definition 3.10. We introduce the concepts of *successor* and *noiseless sequence*. Let (x_k^δ) be a sequence generated by Algorithm 1, and let x_k^δ be an element of this sequence.

- *Successor* of x_k^δ: If $\|F_i(x_k^\delta) - y_i^\delta\| > \tau\delta_i$, a *successor* of x_k^δ is any vector $z \in X$ satisfying $z := x_k^\delta + h$, where either of the following hold:
  (i) h belongs to a pair $(\alpha > \alpha_{\min}, h \in X)$ satisfying (2.5)–(2.6),
  (ii) h belongs to a pair $(\alpha_{\min}, h)$ satisfying (2.5):
  moreover, there exist $\alpha \le \alpha_{\min}$ and $h_\alpha = (F_i'(x_k^\delta)^* F_i'(x_k^\delta) + \alpha I)^{-1} F_i'(x_k^\delta)^*(y_i^\delta - F_i(x_k^\delta))$ satisfying (2.5)–(2.6).
  Otherwise, when $\|F_i(x_k^\delta) - y_i^\delta\| \le \tau\delta_i$, the only *successor* of x_k^δ is the vector $z := x_k^\delta$ itself.
- *Noiseless sequence*: A sequence (x_k) generated by Algorithm 1 in the exact data case is called a *noiseless sequence*.

Some relevant remarks regarding the above definitions follow:
(1) x_{k+1}^δ is a successor of x_k^δ for $k = 0, \dots, k_* - 1$, and x_{k+1} is a successor of x_k for all $k \in \mathbb{N}$.
(2) $x_{k+1} = x_k$ if and only if $F_i(x_k) = y_i$ with $i = [k]$.
(3) In the exact data case $x_{k+1} = x_k + h_k$, where $h_k = \arg\min_h T_{k,\alpha_k}(h)$ with

$$T_{k,\alpha_k}(h) := \|y_i - F_i(x_k) - F_i'(x_k)h\|^2 + \alpha_k\|h\|^2, \tag{3.22}$$

and $\alpha_k > 0$ is defined as in Step 3.2 of Algorithm 1.

Theorem 3.11 (Stability). *Let assumptions* (A1), (A2) *and* (A3) *hold. Let* $(\delta^j) = (\delta_0^j, \dots, \delta_{N-1}^j) \in \mathbb{R}^N$ *be a zero sequence and* $(\mathbf{y}^{\delta^j}) = (y_0^{\delta^j}, \dots, y_{N-1}^{\delta^j}) \in Y^N$ *a corresponding sequence of noisy data satisfying* (1.1). *For each* $j \in \mathbb{N}$, *let* $x_{l+1}^{\delta^j}$ *be a successor of* $x_l^{\delta^j}$ *for* $0 \le l < k_*(\delta^j, \mathbf{y}^{\delta^j})$. *Then there exists a noiseless sequence* (x_l) *such that, for every fixed* $k \in \mathbb{N}$, *there exists a subsequence* $(\delta^{j_m}) \in \mathbb{R}^N$ *(depending on* k*) satisfying*

$$x_l^{\delta^{j_m}} \to x_l \quad \text{as } j_m \to \infty \quad \text{for } l = 0, \dots, k. \tag{3.23}$$

Proof. We use an inductive argument. Since $x_0^\delta = x_0$ for every $\delta \ge 0$, the assertion is clear for $k = 0$. Our main argument consists of recursively choosing subsequences of the original sequence (δ^j). In order to avoid a notational overload, we denote a subsequence of (δ^j) again by (δ^j). Assume by induction that (3.23) holds true for some $k \in \mathbb{N}$, i.e., that there exists a subsequence (δ^j) and $(x_l)_{l=0}^k$ satisfying

$$x_l^{\delta^j} \to x_l \quad \text{as } j \to \infty \quad \text{for } l = 0, \dots, k, \tag{3.24}$$

where x_{l+1} is a successor of x_l for $l = 0, \dots, k-1$.[1] Our goal is to prove the existence of a successor x_{k+1} of x_k and of a subsequence (δ^j) of the current sequence such that $x_{k+1}^{\delta^j} \to x_{k+1}$ as $j \to \infty$, completing the inductive proof. We divide this proof in three steps as follows:

- Step 1. Definition of the element $h_k \in X$ and definition of $x_{k+1} := x_k + h_k$.
- Step 2. Prove that, up to a subsequence, $x_{k+1}^{\delta^j} \to x_{k+1}$ as $j \to \infty$.
- Step 3. Prove that x_{k+1} is a successor of x_k.

Proof of Step 1. Define the sequence $\beta_k^{\delta^j}$ by $\beta_k^{\delta^j} = (\alpha_k^{\delta^j})^{-1}$ if $h_k^{\delta^j} \ne 0$, or $\beta_k^{\delta^j} = 0$ otherwise. Since $k \in \mathbb{N}$ is fixed, it follows from Algorithm 1 that $\beta_k^{\delta^j} \le \alpha_{\min}^{-1}$ for sufficiently large $j \in \mathbb{N}$.[1] Consequently, there exists a subsequence (δ^j) and $\beta_k \in [0, \alpha_{\min}^{-1}]$ such that

$$\beta_k = \lim_{j\to\infty} \beta_k^{\delta^j}. \tag{3.25}$$

If $\beta_k = 0$, define $h_k := 0$. Otherwise, define $\alpha_k := \beta_k^{-1}$ and $h_k := \arg\min_{h\in X} T_{k,\alpha_k}(h)$ (with T_{k,α_k} defined in (3.22)). Now we define and $x_{k+1} := x_k + h_k$.

Proof of Step 2. It is enough to prove that, up to a subsequence, $h_k^{\delta^j} \to h_k$ as $j \to \infty$. From the definition of $h_k^{\delta^j}$ and h_k it follows that

$$\beta_k^{\delta^j}(A_k^{\delta^j})^*(A_k^{\delta^j}h_k^{\delta^j} - b_k^{\delta^j}) + h_k^{\delta^j} = 0 = \beta_k A_k^*(A_k h_k - b_k) + h_k$$

1 Notice that $k_*(\delta^j, \mathbf{y}^{\delta^j}) > k$ for large enough j. Indeed, if this is not the case. the noiseless sequence (x_l) stops after k iterations at a solution of (1.2), and the proof is trivial.

(we use the notation $A_k^{\delta^j} = F_i'(x_k^{\delta^j})$, $b_k^{\delta^j} = y_i^{\delta^j} - F_i(x_k^{\delta^j})$, $A_k = F_i'(x_k)$ and $b_k := y_i - F_i(x_k)$). Consequently,

$$h_k^{\delta^j} = (\beta_k^{\delta^j}(A_k^{\delta^j})^* A_k^{\delta^j} + I)^{-1}(\beta_k^{\delta^j}(A_k^{\delta^j})^* b_k^{\delta^j}) \quad \text{and} \quad h_k = (\beta_k A_k^* A_k + I)^{-1}(\beta_k A_k^* b_k). \tag{3.26}$$

Notice that, from (1.1), (A1) and (3.24) we conclude that $b_k^{\delta^j} \to b_k$ and $A_k^{\delta^j} \to A_k$ as $j \to \infty$. These two facts, together with (3.25) and (3.26), allow us to conclude that $h_k^{\delta^j} \to h_k$ as $j \to \infty$.

Proof of Step 3. We consider two cases:

- First case: $\|F_i(x_k) - y_i\| = 0$. In order to prove that x_{k+1} is a successor of x_k, we have to show that $x_{k+1} = x_k$ (see Definition 3.10). From Step 1. above, we have either $h_k = 0$ or $h_k = \arg\min_{h\in X} T_{k,\alpha_k}(h)$. If $h_k = 0$, then $x_{k+1} = x_k$ and we are done. Otherwise, we have $h_k = (A_k^* A_k + \alpha_k I)^{-1}[A_k^*(y_i - F_i(x_k))] = 0$, and once again we have $x_{k+1} = x_k$.
- Second case: $\|F_i(x_k) - y_i\| > 0$. It follows from (3.24) that $\lim_j \|F_i(x_k^{\delta^j}) - y_i^{\delta^j}\| = \|F_i(x_k) - y_i\| > 0$. Therefore, for sufficiently large $j \in \mathbb{N}$ (say $j > J$), we have $\|F_i(x_k^{\delta^j}) - y_i^{\delta^j}\| > \tau\delta_i^j$ and, consequently, the sequence $(\alpha_k^{\delta^j})_{j>J}$ is defined. Set $\mathcal{J} := \{j > J : \alpha_k^{\delta^j} > \alpha_{\min}\}$. If $\#\mathcal{J} < \infty$, we have, up to a subsequence, $\alpha_k^{\delta^j} = \alpha_{\min}$. Thus, $\alpha_k = \lim_j \alpha_k^{\delta^j} = \alpha_{\min}$. Moreover,

$$(\bar{p} + \eta)\|F_i(x_k^{\delta^j}) - y_i^{\delta^j}\| + (1 + \eta - \bar{p})\delta_i^j \le \|y_i^{\delta^j} - F_i(x_k^{\delta^j}) - A_k h_k^{\delta^j}\|, \quad j > J.$$

  Taking the limit $j \to \infty$, we conclude that $(\bar{p} + \eta)\|F_i(x_k) - y_i\| \le \|y_i - F_i(x_k) - A_k h_k\|$, proving that x_{k+1} is a successor of x_k according to item (ii) of Definition 3.10.

If $\#\mathcal{J} = \infty$, then (2.5) and (2.6) hold true for sufficiently large $j \in \mathbb{N}$. Since $h_k^{\delta^j} \to h_k$ as $j \to \infty$ (see Step 2), we take the limit $j \to \infty$ in (2.5) and (2.6) to conclude that (α_k, h_k) (defined in Step 1) satisfy

$$h_k = (A_k^* A_k + \alpha_k I)^{-1} A_k^*(y_i - F_i(x_k))$$

as well as

$$(\bar{p} + \eta)\|F_i(x_k) - y_i\| \le \|y_i - F_i(x_k) - A_k h_k\| \le (\bar{\bar{p}} + \eta)\|F_i(x_k) - y_i\|,$$

from where we conclude that x_{k+1} is a successor of x_k, according to item (i) of Definition 3.10 (if $\alpha_k > \alpha_{\min}$) or according to item (ii) (if $\alpha_k = \alpha_{\min}$). □

Theorem 3.12 (Semi-convergence). *Let assumptions* (A1), (A2) *and* (A3) *hold. Let* $(\delta^j) = (\delta_0^j, \dots, \delta_{N-1}^j) \in \mathbb{R}^N$ *be a zero sequence and* $(\mathbf{y}^{\delta^j}) = (y_0^{\delta^j}, \dots, y_{N-1}^{\delta^j}) \in Y^N$ *a corresponding sequence of noisy data satisfying* (1.1). *For each* $j \in \mathbb{N}$, *let* $x_{k+1}^{\delta^j}$ *be a successor of* $x_k^{\delta^j}$ *for* $0 \le k < k_*^j = k_*(\delta^j, \mathbf{y}^{\delta^j})$. *Then every subsequence of* $(x_{k_*^j}^{\delta^j})$ *has itself a subsequence converging strongly to a solution of* (1.2).

Proof. Given a subsequence of $(x_{k_*^j}^{\delta^j})$ (which we shall denote again by $(x_{k_*^j}^{\delta^j})$) we consider two distinct cases.

Case I: Assume first that the sequence (k_*^j) is bounded. Then (k_*^j) has a finite accumulation point. Therefore, we can extract a subsequence (δ^{j_m}) such that $k_*^{j_m} = n$ for some $n \in \mathbb{N}$, and all j_m. Theorem 3.11 guarantees the existence of a noiseless sequence (x_k), as well as the existence of a subsequence of $(x_n^{\delta^{j_m}})$ (denoted again by $(x_n^{\delta^{j_m}})$) such that

$$\lim_{m\to\infty} x_{k_*^{j_m}}^{\delta^{j_m}} = \lim_{m\to\infty} x_n^{\delta^{j_m}} = x_n,$$

the nth element of the sequence (x_k). We claim that x_n is a solution of (1.2). Indeed, for each $i \in \{0, \dots, N-1\}$ we have

$$\|F_i(x_n) - y_i\| = \lim_{m\to\infty} \|F_i(x_n^{\delta^{j_m}}) - y_i^{\delta^{j_m}}\| \le \lim_{m\to\infty} \tau\delta_i^{j_m} = 0.$$

Case II: Assume that (k_*^j) is not bounded. Thus, there exists a monotone strictly increasing subsequence, again denoted by (k_*^j). Fix $\varepsilon > 0$ and let (x_k) be a noiseless sequence constructed as in Theorem 3.11. From Theorem 3.9 it follows that $x_k \to x^+$, where x^+ is a solution of (1.2). Then there exists a index $L = L(\varepsilon) \in \mathbb{N}$ such that

$$\|x_k - x^+\| < \frac{\varepsilon}{2}, \quad k \ge L. \tag{3.27}$$

Moreover, since (k_*^j) is unbounded, there exists $J \in \mathbb{N}$ such that $k_*^j \ge L$, for all $j \ge J$. Furthermore, from Corollary 3.2 (see remark after the corollary) it follows that

$$j \ge J \Longrightarrow \|x_{k_*^j}^{\delta^j} - x^+\| \le \|x_L^{\delta^j} - x^+\|.$$

On the other hand, from Theorem 3.11 follows the existence of a subsequence (δ^{j_m}) (depending on L) and the existence of $M \in \mathbb{N}$ such that

$$j_m \geq M \Longrightarrow \|x_L^{\delta^{j_m}} - x_L\| < \frac{\varepsilon}{2}.$$

Thus, for $j_m \geq \max\{J, M\}$ we have

$$\|x_{k_*^{j_m}}^{\delta^{j_m}} - x^+\| \leq \|x_L^{\delta^{j_m}} - x^+\| \leq \|x_L^{\delta^{j_m}} - x_L\| + \|x_L - x^+\| < \varepsilon. \tag{3.28}$$

Notice that the subsequence (δ^{j_m}) depends on ε. We construct an ε–independent subsequence using a diagonal argument: For $\varepsilon = 1$, we can find a subsequence (δ^j) and choose $j_1 \in \mathbb{N}$ such that

$$\|x_{k_*^{j_1}}^{\delta^{j_1}} - x^+\| < 1.$$

Since the current subsequence (δ^j) is also a positive zero-sequence, the same reasoning can be applied again for $\varepsilon = \frac{1}{2}$; then choose $j_2 > j_1$ such that

$$\|x_{k_*^{j_2}}^{\delta^{j_2}} - x^+\| < \frac{1}{2}.$$

Proceeding this way, it is possible to construct a subsequence (δ^{j_n}) satisfying

$$\|x_{k_*^{j_n}}^{\delta^{j_n}} - x^+\| < \frac{1}{n} \quad \text{for all } n \in \mathbb{N}.$$

Consequently, $\|x_{k_*^{j_n}}^{\delta^{j_n}} - x^+\| \to 0$, as $j_n \to \infty$, completing the proof. □

In Theorem 3.12, if we add the assumption in (3.19), we conclude that $(x_{k_*^j}^{\delta^j})$ converges strongly to $x^\dagger$.

4 Numerical experiments

In this section we test the performance of our algorithm in the inverse problem known as Electrical Impedance Tomography (EIT). The first mathematical model of EIT was introduced by A. Calderón in [4] and after that, many variants have been considered. We present a variant, which is regarded as one of the more realistic models, known as the Complete Electrode Model (EIT-CEM), see, e.g., [24].

4.1 Electric Impedance Tomography – Complete Electrode Model

In this procedure, electric currents are injected in the simply connected Lipschitz domain $\Omega \subset \mathbb{R}^2$ via $L \in \mathbb{N}$ electrodes attached to its boundary $\partial\Omega$ and the resulting voltages are measured in the same electrodes with the goal of restoring the electrical conductivity $\gamma : \Omega \to \mathbb{R}$. For the correct translation in an appropriate mathematical model, we suppose that the electrodes $e_1, \ldots, e_L$ are identified with the part of the surface of Ω they contact. Thus, $e_i \subset \partial\Omega$ is open and has positive measure $|e_i| > 0$ for $i = 1, \ldots, L$. Additionally, the electrodes are connected and separated: $\overline{e}_i \cap \overline{e}_j = \varnothing$ for $i \neq j$. In this model, only the density of the electric currents in each electrode e_i, denoted by $I_i \in \mathbb{R}$, are assumed to be known. Further, the electrodes are assumed to be perfect conductors, which means that the electric potential on each electrode e_i is a constant $U_i \in \mathbb{R}$.

The mathematical model which describes the EIT-CEM on the weak formulation is given by the variational equation

$$B((u, U), (v, V)) = \sum_{i=1}^{L} I_i V_i \quad \text{for all } (v, V) \in H, \tag{4.1}$$

where $H := H^1(\Omega) \oplus \mathbb{R}^L_\diamond$ and the bilinear form $B : H \times H \to \mathbb{R}$ is defined by

$$B((u, U), (v, V)) := \int_\Omega \gamma \nabla u \nabla v + \sum_{i=1}^{L} \frac{1}{z_i} \int_{e_i} (u - U_i)(v - V_i)\, \mathrm{dS}. \tag{4.2}$$

Here, $u : \Omega \to \mathbb{R}$ represents the electric potential in Ω, and each number $z_i > 0$ represents the so-called *contact impedance* at the electrode e_i. Further, the vector $U := (U_1, \ldots, U_L)^\top$, which collects the potentials at the electrodes, belongs to $\mathbb{R}^L_\diamond := \{U \in \mathbb{R}^L : \sum_{i=1}^{L} U_i = 0\}$.

In [24] is proved that, if the *current pattern* $I := (I_1, \dots, I_L)^\top$ belongs to the space $\mathbb{R}^L_\diamond$ and the electric conductivity γ belongs to $L^\infty_+(\Omega) := \{\lambda \in L^\infty(\Omega) : \lambda > c > 0\}$, then there exists a unique solution $(u, U) \in H$ satisfying (4.1). This result permits to define the Neumann-to-Dirichlet (NtD) operator $\Lambda_\gamma : \mathbb{R}^L_\diamond \to \mathbb{R}^L_\diamond$, $I \mapsto U$, which is a bounded linear self-adjoint operator. The forward operator associated to EIT-CEM is now defined by the function

$$F : \mathrm{D}(F) \subset L^\infty(\Omega) \to \mathcal{L}(\mathbb{R}^L_\diamond, \mathbb{R}^L_\diamond), \quad \gamma \mapsto \Lambda_\gamma, \tag{4.3}$$

where $\mathrm{D}(F) := L^\infty_+(\Omega)$. Recovering γ from a partial knowledge of Λ_γ is the associated inverse problem we want to solve.

Since the NtD operator Λ_γ is linear, knowledge of the vectors

$$U^j = \Lambda_\gamma I^j, \quad j = 1, \dots, L-1,$$

where $\{I^1, \dots, I^{L-1}\}$ is a basis for $\mathbb{R}^L_\diamond$, is sufficient to determine the NtD operator itself.[2] In practice, one fixes $N \in \mathbb{N}$ current patterns $I^j \in \mathbb{R}^L_\diamond$, $j = 0, \dots, N-1$ (which for notational reasons we put together in a single vector $\mathbb{I} := (I^0, \dots, I^{N-1})^\top \in \mathbb{R}^{NL}$), and reads in the EIT-CEM experiment, a noisy version of the vector $\Gamma_\gamma := (U^0, \dots, U^{N-1})^\top \in \mathbb{R}^{NL}$, where $(u^j, U^j) \in H$ is the unique solution of (4.1) associated with the current pattern $I = I^j$, that is, $U^j := \Lambda_\gamma I^j$, $j = 0, \dots, N-1$. Observe that the noisy versions of U^j belong to the space $\mathbb{R}^L$ but not necessarily to $\mathbb{R}^L_\diamond$.

We now reformulate (4.3) as $F_\mathbb{I} : \mathrm{D}(F) \subset L^\infty(\Omega) \to \mathbb{R}^{NL}$, $\gamma \to \Gamma_\gamma$, and observe that the space $Y = \mathbb{R}^{NL}$ factorizes into the spaces $Y = Y_1 \times \cdots \times Y_N$, with $Y_j := \mathbb{R}^L$, $j = 0, \dots, N-1$, and accordingly, $F_\mathbb{I} = (F_1, \dots, F_N)^\top$ with

$$F_j : \mathrm{D}(F) \subset L^\infty(\Omega) \to \mathbb{R}^L, \quad \gamma \mapsto U^j, \tag{4.4}$$

for $j = 0, \dots, N-1$, which is a more suitable version for the application of a Kaczmarz method.

The operators F_j in (4.4) are Fréchet differentiable, see, e.g., [16, Theorem 4.1]. Let $W^j \in \mathbb{R}^L_\diamond$ be the Fréchet derivative of F_j, evaluated at the vector $\gamma \in \mathrm{int}(\mathrm{D}(F))$ in the direction of $\eta \in L^\infty(\Omega)$, i.e., $W^j := F_j'(\gamma)\eta$. Then W^j is the second element of the pair $(w^j, W^j) \in H$, which is the unique solution of the variational problem

$$B((w^j, W^j), (v, V)) = -\int_\Omega \eta \nabla u^j \nabla v \, \mathrm{d}x \quad \text{for all } (v, V) \in H, \tag{4.5}$$

where $(u^j, U^j) \in H$ is the unique solution of (4.1) with the current pattern $I = I^j$.

We now restrict the searched-for conductivities to a finite-dimensional space. We thus define a triangulation of Ω, $\Upsilon := \{T_i : i = 1, \dots, M\}$, with $M \in \mathbb{N}$ triangles and restrict the solution space X to the finite-dimensional space $V := \mathrm{span}\{\chi_{T_1}, \dots, \chi_{T_M}\} \subset L^\infty(\Omega)$. Accordingly, the domain of definition of F_j is replaced by $V_+ := V \cap L^\infty_+(\Omega) \subset V_2$, where $V_2 := (V, \|\cdot\|_{L^2(\Omega)})$. With this framework, the discrete version of the operator $F_\mathbb{I}$ reads

$$F_\mathbb{I} : V_+ \subset V_2 \to \mathbb{R}^{NL}, \quad \gamma \mapsto (U^1, \dots, U^N). \tag{4.6}$$

Moreover, the operator defined in (4.4) is now given by $F_j : V_+ \subset V_2 \to \mathbb{R}^L$, $\gamma \mapsto U^j$. Summarizing, the inverse problem in (1.1)–(1.2) can be written in the form

$$F_j(\gamma) = U^{j,\delta}, \quad j = 0, \dots, N-1 \tag{4.7}$$

(see Section 4.2.2 for details on the noisy data $U^{j,\delta}$).

Identifying the arbitrary vector $\gamma = \sum_{i=1}^M \alpha_i \chi_{T_i}$ of V_+ with the vector of its coordinates $(\alpha_1, \dots, \alpha_M)^\top \in \mathbb{R}^M$, we observe that the function in (4.7) can now be seen as a non-linear operator from (a subset of) $\mathbb{R}^M$ into $\mathbb{R}^L$. Its derivative evaluated at a vector $\gamma \in \mathrm{int}(V_+)$, $F_j'(\gamma)$, can accordingly be regarded as a matrix (called the *Jacobian* matrix) with dimension $L \times M$. In this framework, the adjoint operator of the derivative is represented by the transposed Jacobian matrix.

In order to calculate the entries of the Jacobian matrix, we use the equation

$$F_j'(\gamma) = (F_j'(\gamma)\chi_{T_1}, \dots, F_j'(\gamma)\chi_{T_M}) \in \mathbb{R}^{L \times M},$$

2 Since Λ_γ is self-adjoint, only $\frac{1}{2}L(L-1)$ potential measurements are actually needed [24].

which demands the computation of M solutions of the variational problem (4.5). However, we employ the strategy explained in [23], which reduces the computational effort to the computation of only L variational problems.

Since analytical solutions of (4.2) and (4.5) are in general not available, the Finite Elements Method (FEM) is used to find approximate solutions. In [17], Lechleiter and Rieder proved that once the triangulation Υ is fixed, there exists a number $L_{\min} \in \mathbb{N}$, depending on Υ, such that if the number of electrodes satisfies $L \geq L_{\min}$, then the Fréchet derivative of the (discrete) forward operator (4.6) is injective and satisfies the weak Tangential Cone Condition (2.2) in a small ball centered on a prefixed arbitrary element $\gamma \in \operatorname{int}(\mathrm{D}(F))$. Further, the same result remains true if the vector $(u, U) \in H$, which is the exact solution of (4.1) is changed by its FEM-approximation. However, to the best of the authors' knowledge, there is no proof of the Tangential Cone Condition for the operators in (4.7).

Finally, in order to obtain more stable reconstructions, we employ a strategy explained in [25], associating each element of the mesh Υ to a positive number called *weight*. As a result, the inner product in space X is changed, and accordingly, the adjoint of the Jacobian matrix $F_j'(\gamma)$ is represented by the transposed matrix, corrected by the given weights [25, Section 4.3].

4.2 Experiments with synthetic data

The computational code used in the implementation of the numerical experiments in this manuscript is written in Python, including the Finite Elements Method (FEM) needed to solve the Partial Differential Equations. The code was implemented with the aid of the open-source platform FEniCS [1] using a Intel® i5 CPU.

In order to test our algorithm using synthetic data, we employ the following procedure:

(1) Fix a refinement Θ of the triangulation Υ.
(2) Define a "solution" γ^+ as a combination of characteristic functions of the elements of Θ.
(3) Generate the (exact) data $U^j := F_j(\gamma^+) \in \mathbb{R}^L$, for $j = 0, \ldots, N-1$ by solving (4.1).[3]
(4) Generate the noisy data $U^{j,\delta}$ by adding by adding artificial noise to U^j.
(5) Recover γ^+ in the original mesh Υ from $U^{j,\delta}$.

The employment of the refined mesh Θ to generate the data U^j is used to avoid inverse crimes. In our experiments, we have chosen the meshes Υ and Θ with 844 and 7568 triangles, respectively. Further, the set Ω is the circle centered at the origin with radius 1.

In all experiments performed in this section, the electrodes have the same size, are equally spaced on the boundary of Ω, and cover 50 % of $\partial\Omega$. We fixed $L = 16$, $N = 15$ and applied the so-called *adjacent current pattern*, which is defined as the set of the vectors $I^j := (0, \ldots, 0, 1, -1, 0, \ldots, 0)^\top$, with 1 in the jth coordinate, -1 in the immediately following one, and zero elsewhere. The contact impedances are the known constants $z_j = 0.025$, for $j = 1, \ldots, L$.

The exact (sought) solution γ^+ is defined by a constant background with conductivity 1 and an inclusion with conductivity 2, i.e.,

$$\gamma^+(x) = \begin{cases} 2 & \text{if } x \in B, \\ 1 & \text{otherwise,} \end{cases}$$

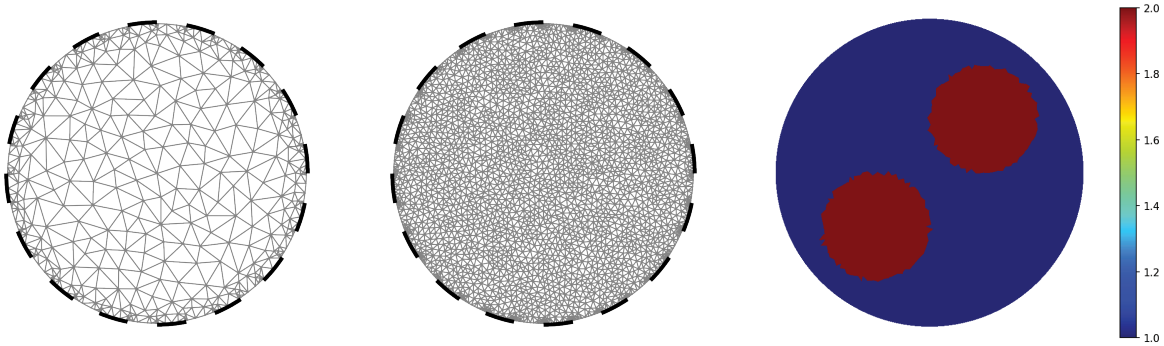
where $B \subset \Omega$ is modeled by two spheric inclusions defined from the fixed mesh. The function γ^+ and the triangulations Θ and Υ are depicted in Figure 1.

In all the tests we use the initial guess $\gamma_0 \equiv 1$, which matches the background. In order compare the results of the experiments, we define the *relative iteration error*

$$E_k = \frac{\|\gamma_k - \gamma^+\|_{L^2(\Omega)}}{\|\gamma^+\|_{L^2(\Omega)}}. \tag{4.8}$$

It is worth noticing that for our choice of γ_0, the initial error is $E_0 = 37.6\,\%$.

3 For a detailed explanation on how to employ the FEM in order to find a solution of (4.1) we refer to [23].



(a) (b) (c)

Figure 1: Experiment with synthetic data: (a) Mesh Υ; (b) Mesh Θ; (c) Sought conductivity γ^+ plotted on the mesh Θ.

4.2.1 Evaluating the Lagrange multipliers

In this subsection we detail the strategy employed for computing the parameters α_k in Step 3.2 of Algorithm 1.

If the norm of the residual $b_k = U^{[k]} - F_{[k]}(\gamma_k)$ is not below the threshold $\tau\delta_{[k]}$ at step k, then it is necessary to compute the derivative $A_k = F'_{[k]}(\gamma_k)$ and to find a parameter $\alpha > 0$ such that the vector h_α (see (2.5)) solution of

$$(A_k^* A_k + \alpha I)h_\alpha = A_k^* b_k \tag{4.9}$$

satisfies (2.6), i.e., $H_k(\alpha) := \|A_k h_\alpha - b_k\| \in [c_k, d_k]$. In order to do so, we implemented an algorithm as follows: Choose constants $0 < a_1 < 1 < a_2$. Choose $\alpha > 0$ and evaluate $H_k(\alpha)$ by applying the Biconjugated Gradient (BCG) method to linear system (4.9). Next, define $\alpha_{\text{old}} = \alpha$ and redefine $\alpha = a_1\alpha_{\text{old}}$ or $\alpha = a_2\alpha_{\text{old}}$ according to $H_k(\alpha) > d_k$ or $H_k(\alpha) < c_k$, respectively. Repeat this process until $H_k(\alpha) \in [c_k, d_k]$.[4]

The computation of the parameters α by the method described above is highly expensive (in the sense of computational effort) in many problems. This is due to the fact that the evaluation of each α_k may require the solution of many linear systems of the form (4.9). However, for the EIT-CEM model, the computation of b_k and A_k are by far the most demanding tasks; once they are computed, the evaluation of the vectors h_α in (4.9), and consequently the evaluation of α_k, do not represent a large amount of computational effort.

4.2.2 Numerical realizations

In our experiments, we compared the rrLMK with two well-known Kaczmarz variations of the Levenberg–Marquardt method with *a priori* choice of the Lagrange multipliers. The first one uses $\alpha_k = \alpha > 0$ constant (stationary LMK or sLMK). The second one uses a geometrical sequence $\alpha_k = r^{\ll k \gg}$,[5] with $0 < r < 1$ (geometric LMK or gLMK).

We add artificially generated noise to the synthetic data U^j with *relative noise* level $\delta > 0$:

$$U^{j,\delta} = U^j + \delta\|U^j\|_2\Delta_j \quad \text{for } j = 0, \ldots, N-1,$$

where the perturbation vectors $\Delta_j \in \mathbb{R}^L$ are uniformly distributed random variables with $\|\Delta_j\|_2 = 1$. Then

$$\frac{\|U^{j,\delta} - U^j\|_2}{\|U^j\|_2} = \delta.$$

We then fix two distinct noise levels: $\delta = 0.1\,\%$ and $\delta = 1\,\%$.

4 After finitely many steps we will find either $H_k(\alpha) \in [c_k, d_k]$ or $[c_k, d_k] \subset [\min\{\alpha, \alpha_{\text{old}}\}, \max\{\alpha, \alpha_{\text{old}}\}]$. In the second case, we apply the bisection method to the function $f(\alpha) = H_k(\alpha) - \frac{1}{2}(c_k + d_k)$ and the interval $[\min\{\alpha, \alpha_{\text{old}}\}, \max\{\alpha, \alpha_{\text{old}}\}]$ until finding $H_k(\alpha) \in [c_k, d_k]$.

5 Here $\ll k \gg$ denotes the largest integer less than or equal to $\frac{k}{N}$, i.e., α_k is constant within each cycle.

Although there are theoretical results [17] that guarantee the existence of a constant η satisfying assumption (A2), it is usually difficult to estimate this constant properly. Aiming to estimate η, we employed a strategy similar to the one in [18]. sLMK was implemented with $\alpha = 10^{-1}$ and $\delta = 0.1\,\%$. Since the constant τ (that depends on η) was not available, we did not check whether or not the discrepancy principle was satisfied in any iteration. Consequently, we did not drop any step during the iterations, but simply performed the fixed number of 180 cycles. During this iteration, we computed the ratios

$$\frac{\|F_{[k]}(y_{k+1}) - F_{[k]}(y_k) - F'_{[k]}(y_k)(y_{k+1} - y_k)\|}{\|F_{[k]}(y_{k+1}) - F_{[k]}(y_k)\|}; \tag{4.10}$$

the largest computed value was multiplied by 2.5 and used as an estimate to η. With this procedure, we found $\eta = 0.4$. The constant τ was then defined by $\tau = \frac{1.3(1+\eta)}{1-\eta}$. The constants $\bar{p}$ and $\bar{\bar{p}}$ (see Algorithm 1) are defined as $[\tau(1-\eta) - (1+\eta)](\tau - 1)^{-1}$ multiplied by 0.01 and 0.99, respectively. The constants a_1 and a_2, needed to compute the Lagrange multipliers (see Section 4.2.1), are set to 0.7 and 1.5, respectively.

The relative iteration error E_k (see (4.8)) at the end of each cycle n (i.e., $k = nN$) is depicted in Figure 2. Pictures (a) and (b) in this figure show the comparison of rrLMK and sLMK, while pictures (c) and (d) show the comparison of rrLMK and gLMK. Pictures (a) and (c) correspond to the relative noise level of $\delta = 0.1\,\%$, while pictures (b) and (d) correspond to $\delta = 1\,\%$,

For the implementation of sLMK we used the constant choices $\alpha_k = 10^{-2}$, $\alpha_k = 10^{-3}$, $\alpha_k = 10^{-4}$ for both levels of noise. For gLMK we used the choices $\alpha_k = 0.98^{\ll k \gg}$, $\alpha_k = 0.95^{\ll k \gg}$, $\alpha_k = 0.9^{\ll k \gg}$ for $\delta = 0.1\,\%$ and the choices $\alpha_k = 0.9^{\ll k \gg}$, $\alpha_k = 0.6^{\ll k \gg}$, $\alpha_k = 0.3^{\ll k \gg}$ for $\delta = 1\,\%$.

Notice that $\alpha_0 = 1$ in gLMK. Thus, in order to be fair in our comparisons, we used $\alpha = 1$ as initial guess to compute the Lagrange multipliers in the first cycle of the rrLMK method (see Section 4.2.1). For the comparison of rrLMK with sLMK we used $\alpha = 10^{-4}$ as initial guess in the first cycle of rrLMK. In the remaining cycles we use as initial guesses for the Lagrange multiplier, the α_k computed in the previous cycle.

In Figure 2 we observe the following facts:

- Comparison of sLMK with rrLMK: These methods perform similarly for both levels of noise. sLMK computes a larger number of cycles than rrLMK for $\alpha_k = 10^{-2}$, it computes a similar number of cycles for $\alpha_k = 10^{-3}$, and it diverges (fails to attain the stop criteria (1.8)) for $\alpha_k = 10^{-4}$.
- Comparison of gLMK with rrLMK for $\delta = 0.1\,\%$: it computes a larger number of steps for $\alpha_k = 0.95^{\ll k \gg}$; it computes an even larger number of steps for $\alpha_k = 0.98^{\ll k \gg}$; it diverges if $\alpha_k = 0.9^{\ll k \gg}$. Comparison of gLMK with rrLMK for $\delta = 1\,\%$: An analog behavior is observed for the choices $\alpha_k = 0.6^{\ll k \gg}$, $\alpha_k = 0.9^{\ll k \gg}$ and $\alpha_k = 0.3^{\ll k \gg}$, respectively.

The number of computed cycles represents an important piece of information to measure the performance of a method, since this is exactly the number of times that the non-linear residual is evaluated (remember that N non-linear residuals are evaluated at each cycle to check which equations satisfy the discrepancy principle). Further, the performance of the method also depends on the number of *active equations* at each cycle (namely, equations that do not satisfy the discrepancy principle at the cycle), since this is the number of derivatives of the forward operator which need to be evaluated at the cycle.

It is worth noticing that the performance of the rrLMK is not significantly affected by the number of linear systems (4.9) solved for each active equation. Thus, the overall number of active equations (denoted here by AcEq) as well as the total number of cycles allow a fair comparison of the performance of these methods.

In Figure 3 the number of active equations in each cycle of the experiments depicted in the first column of Figure 2 is plotted (noise level $\delta = 0.1\,\%$). The left and the right pictures in Figure 3 relate to sLMK and gLMK, respectively.

In Tables 1 and 2 the time elapsed to perform the experiments in Figure 2 is shown. These tables also show the overall number of active equations as well as the number of cycles computed for each method. Table 1 refers to sLMK, while Table 2 refers to gLMK. Notice the direct relation between the AcEq and the curves shown in Figure 3.

In order to compare the reconstructions for the noise level $\delta = 0.1\,\%$, we plot in Figure 4 the last iterate y_{k_*} of each method, namely rrLMK, sLMK and gLMK (correspond to the experiments depicted in the first column of Figure 2). We plot only the best reconstructions obtained using sLMK (with $\alpha_k = 10^{-3}$) and

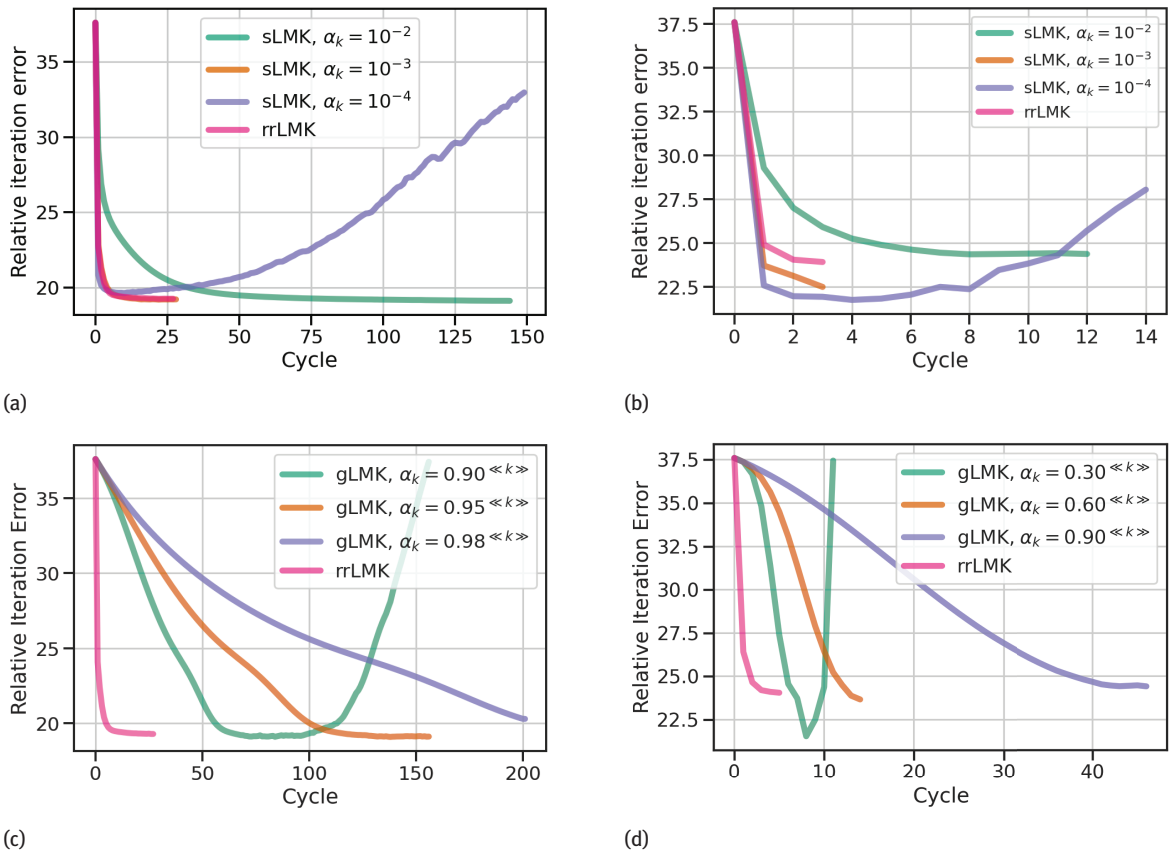


Figure 2: Experiment with synthetic data: Relative iteration error E_k (in %) at the end of each cycle (i.e., $k = nN$): (a) sLMK with $\delta = 0.1\,\%$; (b) sLMK with $\delta = 1\,\%$; (c) gLMK with $\delta = 0.1\,\%$; (d) gLMK with $\delta = 1\,\%$.

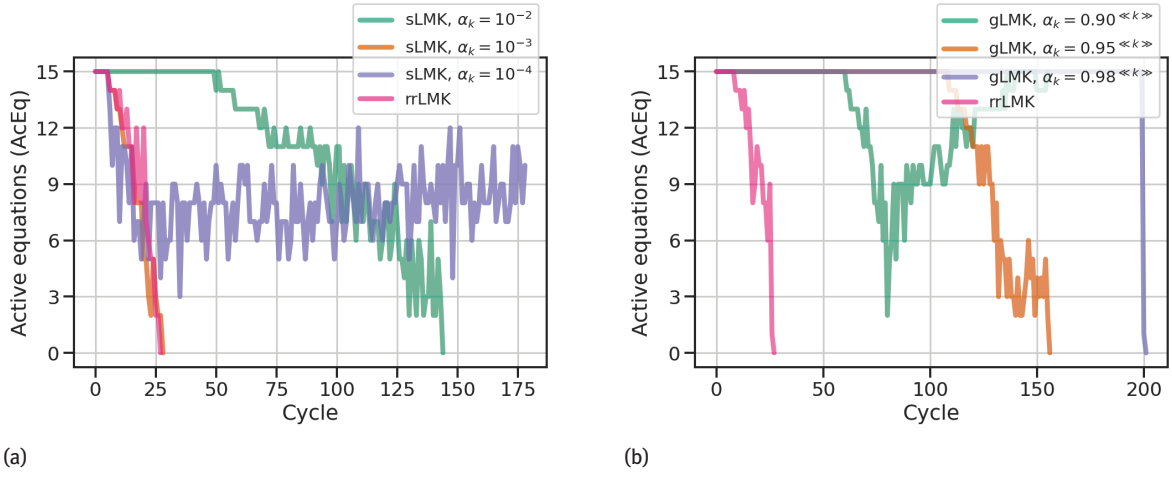


Figure 3: Experiment with synthetic data: Active iterations per cycle with $\delta = 0.1\,\%$. (a) sLMK, (b) gLMK.

gLMK (with $\alpha_k = 0.95^{\ll k \gg}$). Below each picture, the *reconstruction error* at the last iterate E_{k_*} is shown. The reconstructions shown in Figure 4, using different methods, are similar. Apparently, if the parameters are properly chosen, then sLMK and gLMK are capable of delivering reconstructions similar to the one obtained by the rrLMK.

In order to more closely observe how the rrLMK method works, we show in Figure 5 the values of the computed parameters α_k for both noise levels ($\delta = 1\,\%$ and $\delta = 0.1\,\%$) and for all the $N = 15$ equations in (4.7).

Noise level		rrLMK	$\alpha_k \equiv 10^{-4}$	$\alpha_k \equiv 10^{-3}$	$\alpha_k \equiv 10^{-2}$
$\delta = 0.1\,\%$	Time (min)	7.9	Fails	7.69	40.58
	AcEq	294	Fails	274	1617
	Cycles	27	Fails	28	144
$\delta = 1\,\%$	Time (min)	1.2	Fails	1.08	3.28
	AcEq	32	Fails	22	81
	Cycles	3	Fails	3	12

Table 1: Comparison of the performance of rrLMK method and sLMK.

Noise level		rrLMK	$\alpha_k = 0.90^{\ll k \gg}$	$\alpha_k = 0.95^{\ll k \gg}$	$\alpha_k = 0.98^{\ll k \gg}$
$\delta = 0.1\,\%$	Time (min)	8.19	Fails	44.95	62.36
	AcEq	320	Fails	1983	3000
	Cycles	27	Fails	156	200
		rrLMK	$\alpha_k = 0.30^{\ll k \gg}$	$\alpha_k = 0.60^{\ll k \gg}$	$\alpha_k = 0.90^{\ll k \gg}$
$\delta = 1\,\%$	Time (min)	1.8	Fails	4.23	13.14
	AcEq	42	Fails	171	546
	Cycles	5	Fails	14	46

Table 2: Comparison of the performance of rrLMK method and gLMK.

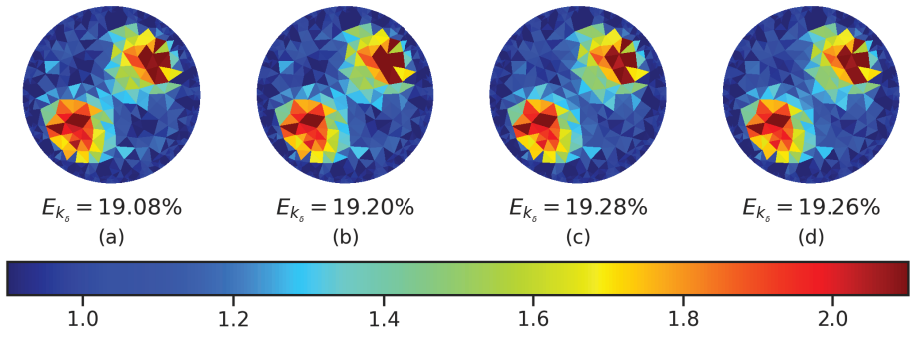


Figure 4: Experiment with synthetic data: Reconstructions γ_{k_*}. From left to the right: sLMK method, gLMK method, rrLMK method with starting guesses $\alpha = 1$ and $\alpha = 10^{-4}$.

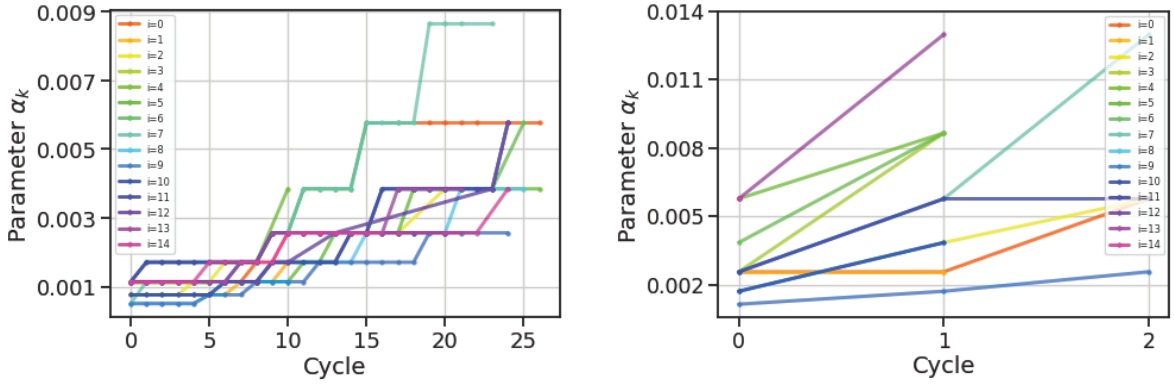


Figure 5: Experiment with synthetic data: Computed values for α_k in the rrLMK method with initial guess $\alpha = 10^{-4}$. (a) $\delta = 0.1\,\%$; (b) $\delta = 1\,\%$.

We conclude this subsection with some final considerations:

(i) In Figure 5 we observe a tendency of the parameters α_k to become larger as the iteration index k grows.

(ii) gLMK performs badly in all considered scenarios, either showing slow convergence or divergence (i.e., not reaching the stop criterion). Further, the "optimal" geometric ratio r ($\alpha_k = r^{\ll k \gg}$) changes according to the noise level δ.

(iii) sLMK performs similarly to the rrLMK if the constant $\alpha_k = \alpha$ is properly chosen. In our numerical tests, the optimal choice of α does not depend on the noise level δ. However, an inadequate choice of α may lead either to slow convergence or to divergence.

(iv) In all numerical tests, the rrLMK method had a better performance than the concurrent sLMK and gLMK methods. The only exception is the test with sLMK with $\alpha_k \equiv 10^{-3}$ (and both levels of noise $\delta = 0.1\,\%$ and $\delta = 1\,\%$), where rrLMK and sLMK performed similarly. It is worth noticing that this situation is significantly different from the observations in [18, 20]. In these references, the gLMK presents relatively good results whenever the constant r in $\alpha_k = r^{\ll k \gg}$ is properly chosen.[6]

4.3 Experiments with real data

In this subsection we reconstruct the solution of inverse problem (4.3) from real data. The experimental framework, as well as the data, were provided by [13, KIT4] from the University of Eastern Finland. All the measurements are made in a prototype, which consists of a cylindrical tank filled with saline water (representing the background) and one or more inclusions made of either metal or plastic. The size, position and shape of the inclusions varies (triangle, cylinder, hollow cylinder, etc.).

The radius of the tank equals 14 cm. $L = 16$ identical equally spaced electrodes made of stainless steel of 2.5 cm width are attached to its boundary and the tank is filled to a height of 7 cm with saline water having the conductivity of 0.03 S/m. Moreover, $N = 15$ adjacent current patterns I^j are employed (see Section 4.2), with amplitude 2mA and frequency of 1kHz.

It is worth noticing that the impedance contacts are not known; the noise levels are not informed, the unit of measurement for the given potentials is not informed. In order to deal with this lack of information, we proceed as follows:

(i) The mesh used to solve the inverse problem is similar to the one used to recover the conductivities in Section 4.2 (see Figure 1). For real data we fix a mesh with 2089 triangles.

(ii) Determining the impedance contacts z_j, $j = 1, \ldots, L$: We assume that they have the same constant value $z > 0$. This value is recovered using the procedure detailed in [25, Definition 4.4] together with the data measured with the tank containing only saline water (without any inclusion). Using this procedure, we obtained the value $z = 2 \cdot 10^{-3}$. This procedure also allow us to estimate the (constant) conductivity of the saline water background, 0.296.[7]

(iii) Determining the noise levels δ_i: We apply the sLMK with $\alpha_k \equiv 0.5$ and initial guess $y_0 \equiv 0.326$ to the data measurements obtained from the tank filled with saline water only. Since the parameter τ, which depends on η and the noise levels δ_j are still unavailable, we do not apply the discrepancy principle neither to terminate the algorithm nor to drop any equation during the iterations, but only run a fixed number of 160 cycles. The resulting *relative residuals*, $\|U^{[k],\delta} - F_{[k]}(y_k)\|/\|U^{[k],\delta}\|$, evaluated for each equation during the iterations, are depicted in Figure 6. We estimate the relative noise levels δ_j, $j = 0, \ldots, N-1$, by making the assumption that the jth noise level is given by the minimum value attained by the residual of the jth equation.[8]

---

6 From considerations (i) and (iv) we infer a possible explanation of why sLMK performs better than the gLMK in our experiments: it seems more appropriate to keep the parameters α_k constant or slowly increasing them during the iterations, instead of decreasing them.

7 This value corresponds to the one informed in [13]. However, it is expressed in another unit and differs from a scaling factor.

8 This assumption is justified by the fact that $\|U^{[k],\delta} - F_{[k]}(y^+)\| \leq \delta_{[k]}$.

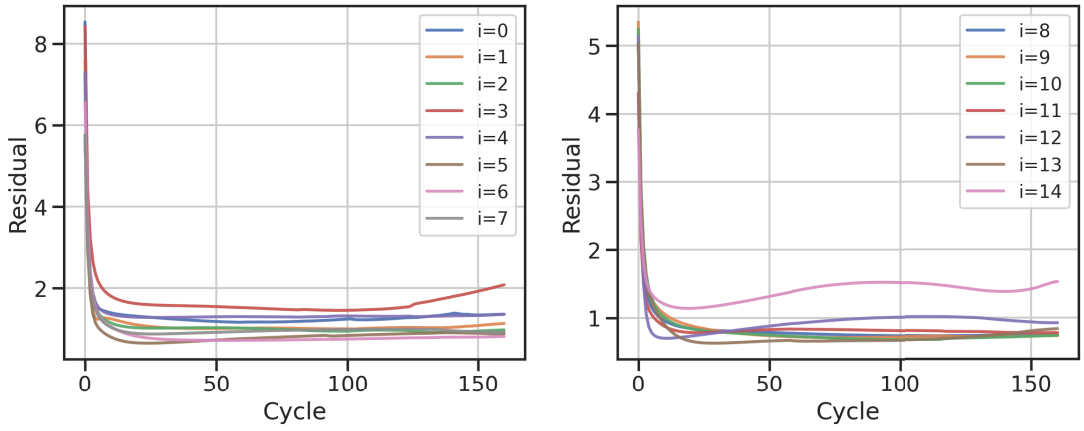


Figure 6: Experiment with real data: Relative residual $r_i(n) := \|F_i(\gamma_{Nn+i}) - U_i\|/\|U_i\|$, $i = 0, \ldots, N-1$, of each equation at the cycle $n = 0, \ldots, 160$, obtained using sLMK with $\alpha_k \equiv 0.5$ and data measured with the tank filled with saline water only.

The constant η in (A2) is estimated using the ratios in (4.10), a procedure analog to the one described in Section 4.2.2. In order to do that, we compute sLMK with $\alpha_k \equiv 0.1$ and the data measurements related to the pictures shown in the first row of Figure 7. This resulted in the value $\eta = 0.25$. The parameter τ was set to $\frac{1.3(1+\eta)}{1-\eta}$. The constants $\bar{p}$ and $\bar{\bar{p}}$ are defined as in the previous experiment.

Methods rrLMK and sLMK are tested[9] to reconstruct three distinct solutions γ_j^+, $j = 1, 2, 3$, all modeled by a saline water background and: γ_1^+ a plastic inclusion with the form of a cylinder; γ_2^+ a metallic inclusion in form of a hollow cylinder; γ_3^+ a metallic and a plastic inclusion in form of a hollow cylinder and a triangle, respectively. The sizes and positions of the inclusions are distinct (see pictures (a), (b) and (c) in Figure 7).

In the experiments shown in Figure 7 we compare reconstructions obtained by the methods rrLMK (2nd row) and sLMK with different constants (3rd row). In order to find the "best" constant for sLMK, we tested several choices for α, namely $\alpha = 10^n$ and $\alpha = 5 \times 10^n$, with $n = -4, -3, -2, -1, 0, 1$. For the choices $n < -1$, the quality of the reconstructions is very poor; for $n = -1$ (i.e., $\alpha = 0.1$ and $\alpha = 0.5$) the reconstructions are close to the sought solutions; for the values $n \geq 0$ the quality of the reconstructions are similar to the choice $n = -1$, however, the corresponding computational effort becomes higher. In Figure 7 (3rd row) we plot the best results obtained using sLMK (with $\alpha = 0.1$ and $\alpha = 0.5$). In Table 3 detailed information on the numerical effort of rrLMK and sLKM methods for these experiments is presented. In Figure 8 we investigate the implementation of the range-relaxed strategy. In what follows we draw some conclusions based on the data presented in Figure 7, Figure 8 and Table 3:

- **[shapes]** What concerns the shapes of the reconstructed inclusions, we observe in Figure 7 that the approximate solutions computed by the sLMK and rrLMK methods have similar quality.
- **[conductivity values]** What concerns the reconstructed conductivity values, we observe in Figure 7 that the reconstructed conductivity of the plastic inclusion is lower than the conductivity of the background (for all methods), while the reconstructed conductivity of the hollow metal inclusion is higher then the conductivity of the background (for all methods). Moreover, in the rrLMK and sLMK methods all reconstructed conductivity values are similar.
- **[computational effort]** The computational effort of the methods under consideration is quite different. In Table 3 we compare the number of computed cycles as well as the overall number of active equations (AcEq) for the experiments conducted in this section (see Figure 7 for a summary). It is worth noticing that, for each method, the related computational effort strongly depends on the solution γ_j^+ to be identi-

9 We also performed numerical tests with the gLMK method. The obtained results are not better (what concerns quality of the reconstruction and numerical effort) then the ones obtained using the sLMK method. For this reason these results are not described in this section.

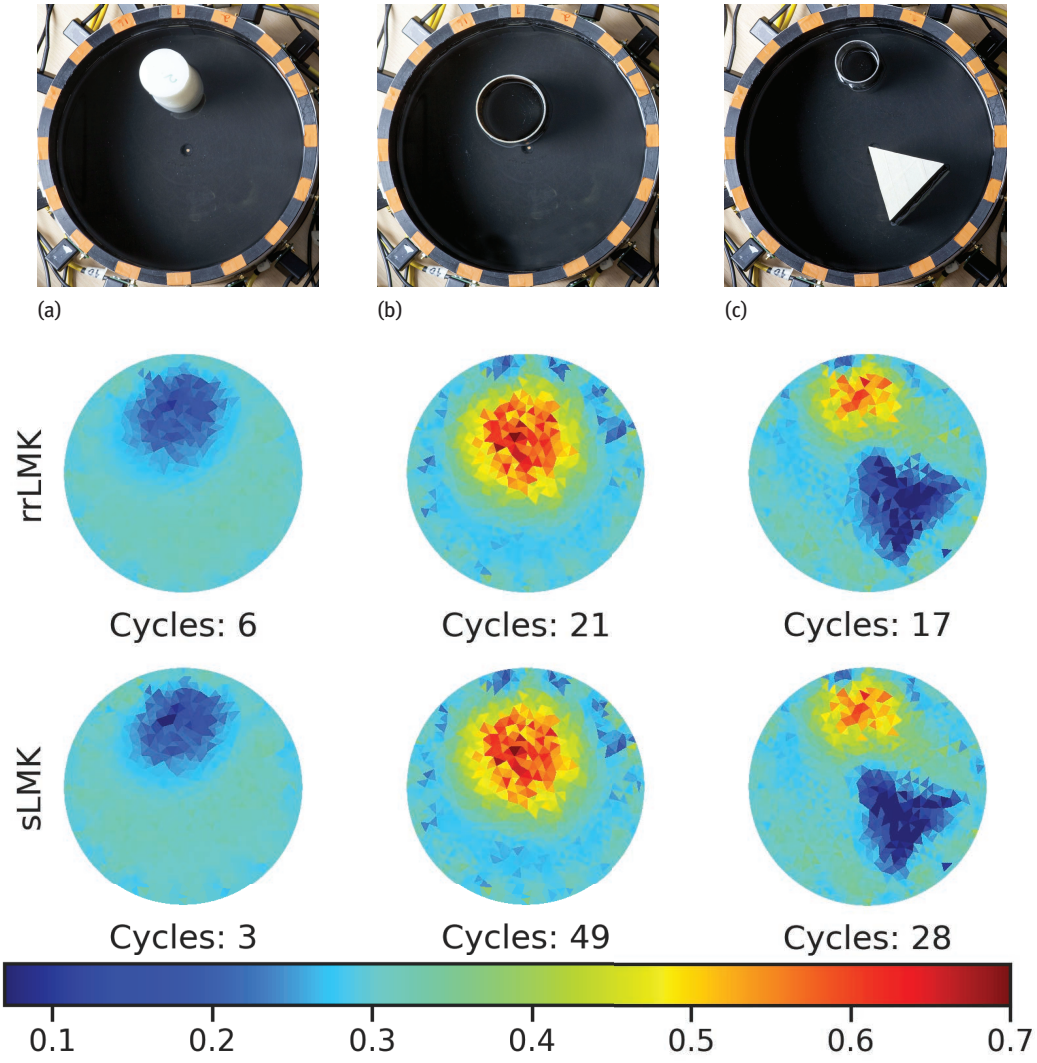
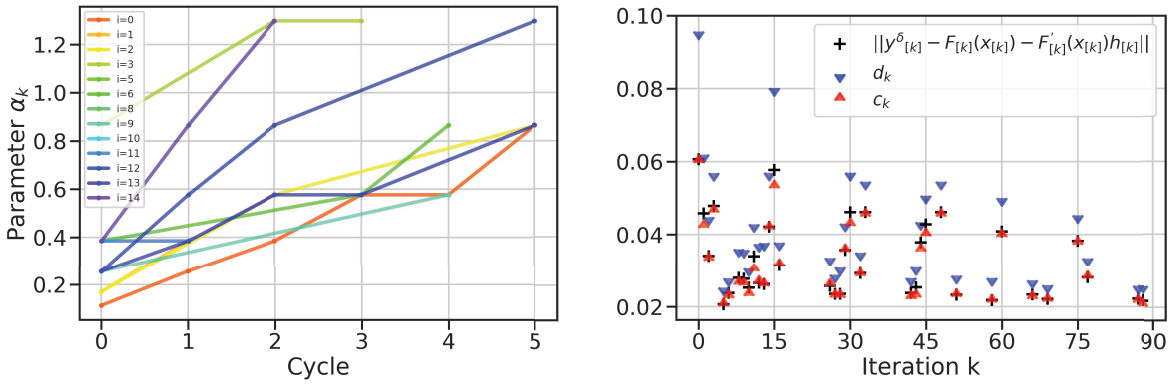


Figure 7: Experiment with real data: Reconstructions γ_{k_*}. First row: real parameters. Second row: rrLMK method. Third row: From left to the right sLMK with constant $\alpha = 0.1$, $\alpha = 0.5$ and $\alpha = 0.5$.

Solution		γ_1^+	γ_2^+	γ_3^+
Time (min)	rrLMK	3.03	7.61	7.21
	sLMK	0.99	11.78	7.88
AcEq	rrLMK	36	158	143
	sLMK	20	308	207
Cycles	rrLMK	6	21	17
	sLMK	3	49	28

Table 3: Comparison of the computational effort of the sLMK and the rrLMK methods when applied to reconstruct the pictures shown in Figure 7.



(a) (b)

Figure 8: Experiment with real data: implementation of the rrLMK method when applied to reconstruct the image (a) in Figure 7. (a) Evolution of the parameters α_k. (b) Linearized residual and the numbers c_k and d_k in each iteration k.

fied. In Table 3 we observe that, if the constant $\alpha_k = \alpha$ is chosen properly, the sLMK becomes competitive with rrLMK from the computational point of view.

- **[range-relaxed strategy]** Figure 8 provides more information concerning the iteration of rrLMK to reconstruct the solution y_1^+ (see Figure 7 (a)). In Figure 8 (a) we presents the evolution of the parameters α_k. Similarly to the results observed in Section 4.2, the parameters α_k tend to increase with the iteration index k. In Figure 8 (b) we display the values of the linearized residuals $\|F'_{[k]}(y_k)(y_{k+1} - y_k) - (U^{[k],\delta} - F_{[k]}(y_k))\|$ and the constants c_k and d_k (see Algorithm 1) at each step. Notice that, as required,

$$\|F'_{[k]}(y_k)(y_{k+1} - y_k) - (U^{[k],\delta} - F_{[k]}(y_k))\| \in [c_k, d_k] \quad \text{for } k = 0, \ldots, k_*.$$

5 Conclusions

We investigate LMK-type methods for computing stable approximate solutions to systems of non-linear ill-posed operator equations. The main contribution of this article is to extend the strategy for choosing a sequence of Lagrange multipliers in [18] (we propose a different range) in order to couple the rrLM method [18] with the Kaczmarz strategy. This modification allow us to prove convergence for exact data (Section 3) using a technique different from the one applied in [18]. Moreover, we also prove stability and semi-convergence results.

The range-relaxed LMK (rrLMK) method is advantageous when compared with other *a posteriori* strategies for computing the Lagrange multipliers, since it allows each of the multipliers to belong to a non-degenerate interval. Consequently, the actual computation of the Lagrange multipliers (satisfying the theoretical requirements needed for the convergence analysis) is simplified.

It is worth noticing that LM-type methods based on *a priori* choices for the multipliers (e.g., gLMK and sLMK) require additional information on the sequence of multipliers in order to be efficiently implemented. This information is usually obtained by trial and error. In the rrLMK method, on the other hand, the multipliers are computed/adjusted during the iteration and no additional *a priori* information on the behavior of the multipliers is required.

An algorithmic implementation of the rrLMK method (Algorithm 1) is discussed, and it is tested for solving a well known ill-posed problem (Electric Impedance Tomography – Complete Electrode Model [24]) using both synthetic and real data. For this particular mathematical model, the evaluation of the Lagrange multiplier using the range-relaxed strategy does not represent a large computational cost (see Section 4.2.1). Consequently, the cost of one rrLMK iteration is comparable with the cost of one iteration of the sLMK and gLMK methods.

The numerical experiments in Section 4 include real data sets using the framework proposed in [13]. All the measurements are made in a prototype, which consists of a cylindrical tank (of 14 cm radius) filled with saline water (representing the background), 16 identical stainless steel electrodes and one or more inclusions made of either metal or plastic. Our experiments show (in both synthetic and real data cases) that the rrLMK method is more efficient than the other tested Kaczmarz-type methods, namely the gLMK method (where the Lagrange multipliers are computed *a priori* in geometric progression) and the sLMK method (with constant Lagrange multipliers).

Funding: Joel C. Rabelo acknowledges support from the Federal University of Piaui. Antonio Leitão acknowledges support from the research agency CNPq (grant 311087/2017-5), and from the AvH Foundation. Eduardo Hafemann acknowledges support from CNPq.

References

[1] M. S. Alnæs, J. Blechta, J. Hake, A. Johansson, B. Kehlet, A. Logg, C. Richardson, J. Ring, M. E. Rognes and G. N. Wells, The fenics project version 1.5., *Arch. Numer. Softw.* **3** (2015), no. 100, 9–23.

[2] 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** (2010), no. 3, 335–350.

[3] M. Burger and B. Kaltenbacher, Regularizing Newton–Kaczmarz methods for nonlinear ill-posed problems, *SIAM J. Numer. Anal.* **44** (2006), no. 1, 153–182.

[4] A.-P. Calderón, On an inverse boundary value problem, in: *Seminar on Numerical Analysis and its Applications to Continuum Physics* (Rio de Janeiro 1980), Brazilian Mathematical Society, Rio de Janeiro (1980), 65–73.

[5] A. De Cezaro, M. Haltmeier, A. Leitão and O. Scherzer, On steepest-descent-Kaczmarz methods for regularizing systems of nonlinear ill-posed equations, *Appl. Math. Comput.* **202** (2008), no. 2, 596–607.

[6] H. W. Engl, M. Hanke and A. Neubauer, *Regularization of Inverse Problems*, Math. Appl. 375, Kluwer Academic, Dordrecht, 1996.

[7] R. Filippozzi, J. C. Rabelo, R. Boiger and A. Leitão, A range-relaxed criteria for choosing the Lagrange multipliers in the iterated Tikhonov Kaczmarz method for solving systems of linear ill-posed equations, *Inverse Problems* **37** (2021), no. 4, Article ID 045005.

[8] M. Haltmeier, R. Kowar, A. Leitão and O. Scherzer, Kaczmarz methods for regularizing nonlinear ill-posed equations. II. Applications, *Inverse Probl. Imaging* **1** (2007), no. 3, 507–523.

[9] M. Haltmeier, A. Leitão and E. Resmerita, On regularization methods of EM-Kaczmarz type, *Inverse Problems* **25** (2009), no. 7, Article ID 075008.

[10] M. Haltmeier, A. Leitão and O. Scherzer, Kaczmarz methods for regularizing nonlinear ill-posed equations. I. Convergence analysis, *Inverse Probl. Imaging* **1** (2007), no. 2, 289–298.

[11] M. Hanke, A regularizing Levenberg–Marquardt scheme, with applications to inverse groundwater filtration problems, *Inverse Problems* **13** (1997), no. 1, 79–95.

[12] M. Hanke, A. Neubauer and O. Scherzer, A convergence analysis of the Landweber iteration for nonlinear ill-posed problems, *Numer. Math.* **72** (1995), no. 1, 21–37.

[13] A. Hauptmann, V. Kolehmainen, N. M. Mach, T. Savolainen, A. Seppänen and S. Siltanen, Open 2d electrical impedance tomography data archive, 2017.

[14] S. Kaczmarz, Approximate solution of systems of linear equations, *Internat. J. Control* **57** (1993), no. 6, 1269–1271.

[15] B. Kaltenbacher, A. Neubauer and O. Scherzer, *Iterative Regularization Methods for Nonlinear Ill-Posed Problems*, Radon Ser. Comput. Appl. Math. 6, Walter de Gruyter, Berlin, 2008.

[16] A. Lechleiter and A. Rieder, Newton regularizations for impedance tomography: A numerical study, *Inverse Problems* **22** (2006), no. 6, 1967–1987.

[17] A. Lechleiter and A. Rieder, Newton regularizations for impedance tomography: Convergence by local injectivity, *Inverse Problems* **24** (2008), no. 6, Article ID 065009.

[18] A. Leitão, F. Margotti and B. F. Svaiter, Range-relaxed criteria for choosing the Lagrange multipliers in the Levenberg–Marquardt method, *IMA J. Numer. Anal.* **41** (2021), no. 4, 2962–2989.

[19] K. Levenberg, A method for the solution of certain non-linear problems in least squares, *Quart. Appl. Math.* **2** (1944), 164–168.

[20] M. P. Machado, F. Margotti and A. Leitão, On the choice of Lagrange multipliers in the iterated Tikhonov method for linear ill-posed equations in Banach spaces, *Inverse Probl. Sci. Eng.* **28** (2020), no. 6, 796–826.

[21] D. W. Marquardt, An algorithm for least-squares estimation of nonlinear parameters, *J. Soc. Indust. Appl. Math.* **11** (1963), 431–441.

[22] F. Natterer, *The Mathematics of Computerized Tomography*, Class. Appl. Math. 32, Society for Industrial and Applied Mathematics, Philadelphia, 2001.

[23] N. Polydorides and W. R. B. Lionheart, A Matlab toolkit for three-dimensional electrical impedance tomography: A contribution to the electrical impedance and diffuse optical reconstruction software project, *Measurement Sci. Technol.* **13** (2002), no. 12, 1871–1873.

[24] E. Somersalo, M. Cheney and D. Isaacson, Existence and uniqueness for electrode models for electric current computed tomography, *SIAM J. Appl. Math.* **52** (1992), no. 4, 1023–1040.

[25] R. Winkler, *A model-aware inexact Newton scheme for electrical impedance tomography*, Ph.D. Thesis, Karlsruher Institut für Technologie, Karlsruhe, 2016.