CHALMERS | UNIVERSITY OF GOTHENBURG

PREPRINT 2012:4

Relaxation Property of the Adaptivity Technique for Some Ill-Posed Problems

LARISA BEILINA
MICHAEL V. KLIBANOV

Department of Mathematical Sciences
Division of Mathematics
CHALMERS UNIVERSITY OF TECHNOLOGY
UNIVERSITY OF GOTHENBURG
Gothenburg Sweden 2012

Preprint 2012:4

Relaxation Property of the Adaptivity Technique for Some Ill-Posed Problems

Larisa Beilina and Michael V. Klibanov

Department of Mathematical Sciences
Division of Mathematics
Chalmers University of Technology and University of Gothenburg
SE-412 96 Gothenburg, Sweden

Gothenburg, April 2012

Preprint 2012:4
ISSN 1652-9715

Matematiska vetenskaper
Göteborg 2012

RELAXATION PROPERTY OF THE ADAPTIVITY TECHNIQUE FOR SOME ILL-POSED PROBLEMS

LARISA BEILINA * AND MICHAEL V. KLIBANOV †

Abstract. Adaptive Finite Element Method (adaptivity) is known to be an effective numerical tool for some ill-posed problems. The key advantage of the adaptivity is the image improvement with local mesh refinements. A rigorous proof of this property is the *central* part of this paper. In terms of Coefficient Inverse Problems with single measurement data, the authors consider the adaptivity as the second stage of a two-stage numerical procedure. The first stage delivers a good approximation of the exact coefficient without an advanced knowledge of a small neighborhood of that coefficient. This is a necessary element for the adaptivity to start iterations from. Numerical results for the two-stage procedure are presented for both computationally simulated and experimental data.

1. Introduction. This paper summarizes recent results of the authors on the Adaptive Finite Element Method (adaptivity) for Ill-Posed problems, see [1, 6, 10, 11, 12, 13, 14, 15]. Results, which are formulated in the setting of Functional Analysis, are applied to a Coefficient Inverse Problem (CIP) for a hyperbolic PDE. Both formulations and proofs of almost all theorems are modified here, compared with above publications. The *central feature* of the adaptivity is the relaxation property. In short, relaxation means that the accuracy of the computed solutions improves with refinements of meshes of finite elements. The relaxation property for the adaptivity for CIPs was observed numerically in many publications, see, e.g. [1, 6, 7, 8, 9, 10, 12, 13, 14]. Analytically it was established for the first time in [15]. Here we present a simpler proof of this property (Theorem 5.2). Since this simpler proof was published only in the book [11] (Theorem 4.9.3), it makes sense to have a journal publication of this result. It is shown in section 8 (Remark 8.2) that the relaxation property helps to work out the stopping criterion for mesh refinements. Theorems 5.3 and 5.4 as well as the numerical example of Test 1 were not published before.

The adaptivity for a CIP for a hyperbolic PDE was first proposed in 2001 in [6]. Also, in 2001 a similar idea was proposed in [4], although an example of a CIP was not considered in [4]. In both these first publications the so-called " Galerkin orthogonality principle" was used quite essentially. The adaptivity was developed further in a number of publications, where it was applied both to CIPs [3, 6, 7, 8, 9, 10] and to the parameter identification problems, which are different from CIPs, see, e.g. [23, 24, 25]. In the recent publication [37] the adaptivity was applied to the Cauchy problem for the Laplace equation. Unlike other works, both lower and upper error estimates were obtained in [37]. In another recent publication [33] an adaptive finite element method for the solution of a Fredholm integral equation of the first kind was presented and *a posteriori* error estimates for both the Tikhonov functional and the regularized solution were derived.

The essence of the adaptivity consists in the minimization of the Tikhonov functional on a sequence of locally refined meshes. Only standard piecewise finite elements are used in this paper. It is important that due to local rather than global mesh refinements, the total number of finite elements is rather moderate. If this number would be very large, then the corresponding space of finite elements would effectively behave as an infinitely dimensional one. However, in the case of a moderate number of finite elements, this space effectively behaves as a finite dimensional one. This is the underlying reason why our *main interest* in this paper is in constructing the theory of the adaptivity only in a *finite dimensional space* of finite elements. Since all norms in finite dimensional spaces are equivalent, then we use the same norm in the Tikhonov regularization term as the one in the original space (except of subsection 2.1). This is obviously more convenient for both analysis and numerical studies than the standard case of a stronger norm in this term [2, 11, 26, 42, 43]. Numerical results of the current and previous publications confirm the validity of this approach. Note that although the finite dimensional version of the original ill-posed problem might be well

* Department of Mathematical Sciences, Chalmers University of Technology and Gothenburg University, SE-42196 Gothenburg, Sweden, (larisa@chalmers.se)

† Department of Mathematics and Statistics University of North Carolina at Charlotte, Charlotte, NC 28223, USA, (mklibanv@uncc.edu)

posed, at least formally, in the actuality it inherits the ill-posedness at certain extent. Thus, the use of the regularization term is still important for the stabilization.

Recall that a minimizer of the Tikhonov functional, if it exists, is called *regularized solution* of the corresponding equation [2, 11, 20, 26, 42, 43]. In the case of a nonlinear problem, such as, e.g. a CIP, many regularized solutions might exists or none. Furthermore, even if a regularized solution exists and is unique, it is unclear how to practically find it, unless a good first guess about the true solution is available. This is because of the well known phenomenon of multiple local minima and ravines of the Tikhonov functional.

In the past several years the authors have addressed five important questions about the adaptivity, which were not addressed before. These five questions are discussed in the current paper:

1. It was established in [32] that if the operator of an ill-posed problem is one-to-one, then a regularized solution is indeed closer to the exact one than the first guess for a *single* pair $(\delta, \alpha(\delta))$. Here $\delta > 0$ is the level of the error in the data and $\alpha(\delta)$ is the regularization parameter, see section 2. Note that in the classical regularization theory convergence of regularized solutions to the exact one is claimed only in the limiting case when $(\delta, \alpha(\delta)) \to (0,0)$ [11, 20, 42, 43]. However, in practical computations one always works only with a single pair $(\delta, \alpha(\delta))$.

2. The local strong convexity of the Tikhonov functional was established under the condition that the originating operator F has the first Fréchet derivative F' satisfying the Lipschitz condition. First, it was established in a small neighborhood of a regularized solution [15]. In the current paper the proof of [15] is modified to obtain the local strong convexity in a small neighborhood of the exact solution, which is more useful. In the finite dimensional case, which is our main interest because of finite elements (see above), it is shown below that the local strong convexity guarantees both existence and uniqueness of the regularized solution. In the previous publication [40] the local strong convexity was established for the case when the originating operator F has the second continuous Fréchet derivative and the source representation condition is in place. We do not use the source representation condition. In [23] the local strong convexity of the Tikhonov functional was assumed rather than proved.

3. Estimates of distances between regularized solutions and ones obtained after mesh refinements were obtained. In the adaptivity these estimates are usually called "*a posteriori* error estimates" (Theorems 5.1, 6.4 and 6.5 below). Unlike this, in the past publications about the adaptivity for ill-posed problems, such estimates were obtained only for some functionals rather than for solutions themselves, see, e.g. [1, 4, 6, 7, 9, 10, 12, 13, 24, 25].

4. The relaxation property was established [11, 15]. In particular, it follows from this property that the accuracy of the reconstruction of the exact solution improves in the mesh refinement process, provided that maximal grid step sizes of those mesh refinements are properly chosen, see Theorem 5.3, Remark 5.1 and Theorem 6.5 below. To prove the relaxation property, a new framework of Functional Analysis for the adaptivity technique was introduced (section 4).

5. A two stage numerical procedure was developed for some CIPs for a hyperbolic PDE [11, 12, 13, 14]. On the first stage, the approximately globally convergent method delivers a good approximation for the exact solution. On the second stage, the adaptivity uses this approximation as a starting point for a refinement.

The reason why we need the first stage is that the adaptivity works only in a small neighborhood of the exact solution. The question on how to actually obtain a point in that neighborhood is left outside of this theory. This question was recently addressed by the authors for CIPs for a hyperbolic PDE with single measurement data, see, e.g. [11, 12, 13, 14, 15, 31, 34, 35] and further referenced cited there. It is well known that addressing this question for CIPs is an *enormously challenging* goal. This is because of two factors combined: nonlinearity and ill-posedness of CIPs. Therefore, one inevitably faces a tough dilemma: either ignore this question, or try to address it, but within the framework of a certain reasonable approximate mathematical model. As to that model, see sections 2.9 and 6.6.2 of [11] as well as [34, 35] for it. We call corresponding algorithms "approximately globally convergent". It was rigorously established, within the framework of that approximate model, that the numerical method of these publications indeed delivers a point in a small neighborhood of the exact solution of the corresponding CIP without any advanced knowledge of that neighborhood. The validity of that approximate mathematical model was verified via a six-step procedure, see section 1.1.2 of [11] as well as [34, 35] for this procedure. Basically the verification amounts to the numerical testing of the corresponding algorithm on computationally simulated data as well as, most importantly, on *blind* experimental data, see [31, 35] and also Chapter 5 and section 6.9 of the book [11] for blind data studies. "Blind" means that the solution was unknown in advance. Since the blind experimental data case is unbiased, then the verification for blind data is the most persuasive one.

In the sections 2-5 we use the apparatus of the Functional Analysis to address above items 1-4 for rather general ill-posed problems. In section 6 we deduce from sections 2-5 some results for a CIP for a hyperbolic PDE. In section 7 we present mesh refinement recommendations. In section 8 we present numerical results, including ones for real experimental data. In numerical studies of this paper we use the above mentioned two-stage numerical procedure.

2. Minimizing Sequence and a Regularized Solution Versus the First Guess.

2.1. The infinitely dimensional case. Let B, B_1, B_2 be three Banah spaces. We denote norms in these spaces respectively as $\|\cdot\|, \|\cdot\|_1, \|\cdot\|_2$. As it is conventional in the theory of Ill-Posed problems, we assume that $B_1 \subseteq B, \|x\| \leq C\|x\|_1, \forall x \in B_1$ and $\overline{B}_1 = B$, $C = const. > 0$, and the closure $\overline{B}_1$ is in the norm $\|\cdot\|$. Furthermore, we assume that any bounded set in B_1 is a compact set in B. Let $G \subseteq B_1$ be a set and $\overline{G}$ be its closure in the norm $\|\cdot\|$. Let $F : \overline{G} \to B_2$ be a continuous operator in terms of norms $\|\cdot\|, \|\cdot\|_2$. Consider the equation

$$F(x) = y, x \in G. \tag{2.1}$$

As it is usually done in the regularization theory [2, 11, 20, 26, 42, 43], we assume that the right hand side of equation (2.1) is given with a small error $\delta \in (0,1)$. We also assume that there exists an "ideal" exact solution x^* of (2.1) with the "ideal" exact data y^* (in principle, there might be several exact solutions). Thus, we assume that

$$F(x^*) = y^*, x^* \in G, \|y - y^*\|_2 \leq \delta. \tag{2.2}$$

Let $x_0 \in B_1$ be a first guess for the exact solution x^*. Usually one assumes that x_0 is located in a small neighborhood of x^*. Consider the Tikhonov functional

$$M_\alpha(x) = \frac{1}{2}\|F(x) - y\|_2^2 + \frac{\alpha}{2}\|x - x_0\|_1^2, x, x_0 \in G, \tag{2.3}$$

where $\alpha \in (0,1)$ is the regularization parameter. The second term in the right hand side of (2.3) is called "the Tikhonov regularization term". Let

$$m_\alpha = \inf_G M_\alpha(x). \tag{2.4}$$

Hence, there exists a minimizing sequence $\{x_n^\alpha\}_{n=1}^\infty \subset G$ such that

$$\lim_{n\to\infty} M_\alpha(x_n^\alpha) = m_\alpha. \tag{2.5}$$

Theorem 2.1 positively addresses the following question: *Let both numbers $\delta, \alpha = \alpha(\delta)$ be fixed. Does the minimizing sequence in (2.5) deliver a better approximation for the exact solution x^* than the first guess x_0?* Since $\|x_0 - x^*\|_1$ is usually assumed to be small, we assume for the sake of defineteness that $\|x_0 - x^*\|_1 \leq 1$.

Theorem 2.1. *Let B, B_1, B_2 be Banah spaces, $G \subset B_1$ be a convex open set and $F : \overline{G} \to B_2$ be a one-to-one continuous operator in terms of norms $\|\cdot\|, \|\cdot\|_2$. Let $\alpha = \alpha(\delta) = \delta^{2\mu}, \mu = const. \in (0,1/2)$ and conditions (2.4), (2.5) are in place. Let $\|x_0 - x^*\|_1 \leq 1$ and $\xi \in (0,1)$ be an arbitrary number. Assume first that $x_0 \neq x^*$. Then there exists a sufficiently small number $\delta_0 = \delta_0(F, \mu, \xi) \in (0,1)$ such that for any $\delta \in (0, \delta_0)$ there exists an integer $N = N(\delta, F) \geq 1$ such that for*

$$\left\|x_n^{\alpha(\delta)} - x^*\right\| \leq \xi\|x_0 - x^*\| \leq C\xi\|x_0 - x^*\|_1, \forall n \geq N. \tag{2.6}$$

If $x_0 = x^*$*, then (2.6) should be replaced with*

$$\left\| x_n^{\alpha(\delta)} - x^* \right\| \leq \xi. \tag{2.7}$$

In particular, if $\delta = 0$*, then* δ_0 *should be replaced with a sufficiently small number* $\alpha_0 \in (0,1)$ *and* $\delta \in (0, \delta_0)$ *should be replaced with* $\alpha \in (0, \alpha_0)$*.*

Proof. By (2.2)-(2.4)

$$m_\alpha \leq M_\alpha(x^*) < \delta^2 + \alpha \|x_0 - x^*\|_1^2. \tag{2.8}$$

Hence, there exists an integer $N = N(\delta, F) \geq 1$ such that $M_\alpha\left(x_n^{\alpha(\delta)}\right) < \delta^2 + \alpha \|x_0 - x^*\|_1, \forall n \geq N$. Hence, by (2.2)

$$\left\| x_n^{\alpha(\delta)} \right\|_1 \leq \sqrt{2}\left(\delta^{2(1-\mu)} + \|x_0 - x^*\|_1^2\right)^{1/2} + \|x_0\|_1, \forall n \geq N(\delta, F). \tag{2.9}$$

Consider the set $P(\delta, x_0)$ defined as

$$P(\delta, x_0) = \left\{ x \in G : \|x\|_1 \leq \sqrt{2}\left(\delta^{2(1-\mu)} + \|x_0 - x^*\|_1^2\right)^{1/2} + \|x_0\|_1 \right\}. \tag{2.10}$$

Let $\overline{P} := \overline{P}(\delta, x_0)$ be its closure in terms of the norm $\|\cdot\|$. Hence, $\overline{P}(\delta, x_0) \subseteq \overline{G}$. Since the set $P(\delta, x_0)$ is bounded in terms of the norm $\|\cdot\|_1$, then $\overline{P}$ is a closed compact set in the space B. Consider the range $F\left(\overline{P}\right) \subset B_2$ of the operator F on the set $\overline{P}$. Since the operator $F : \overline{G} \to B_2$ is continuous in terms of norms $\|\cdot\|, \|\cdot\|_2$, then $F\left(\overline{P}\right)$ is a closed compact set in B_2. Furthermore, since F is one-to-one, then by the foundational theorem of Tikhonov [11, 42, 43] the inverse operator $F^{-1} : F\left(\overline{P}\right) \to \overline{P}$ is continuous. Therefore, there exists the modulus of the continuity of the operator F^{-1} on the set $F\left(\overline{P}\right)$. This means that there exists a function $\omega_F(z), z \in (0,1)$ such that

$$\omega_F(z) \geq 0, \omega_F(z_1) \leq \omega_F(z_2) \text{ if } z_1 \leq z_2, \lim_{z \to 0^+} \omega_F(z) = 0, \tag{2.11}$$

$$\|x_1 - x_2\| \leq \omega_F\left(\|F(x_1) - F(x_2)\|_2\right), \forall x_1, x_2 \in \overline{P}. \tag{2.12}$$

Using (2.2) and (2.8), we obtain for $n \geq N(\delta, F)$

$$\left\| F\left(x_n^{\alpha(\delta)}\right) - F(x^*) \right\|_2 = \left\| F\left(x_n^{\alpha(\delta)}\right) - y + y - F(x^*) \right\|_2 \tag{2.13}$$

$$\leq \left\| F\left(x_n^{\alpha(\delta)}\right) - y \right\|_2 + \|y - y^*\|_2 \leq \left[2M_\alpha\left(x_n^{\alpha(\delta)}\right)\right]^{1/2} + \delta \leq \left(\delta^2 + \delta^{2\mu}\|x_0 - x^*\|_1^2\right)^{1/2} + \delta.$$

By (2.8) and (2.10) $x^* \in \overline{P}$. Therefore, (2.12) and (2.13) imply that there exists a sufficiently small number $\widetilde{\delta} = \widetilde{\delta}(F, \mu) \in (0,1)$ such that for all $\delta \in \left(0, \widetilde{\delta}\right)$ we have $\left(\delta^2 + \delta^{2\mu}\|x_0 - x^*\|_1^2\right)^{1/2} + \delta < 1$ and

$$\left\| x_n^{\alpha(\delta)} - x^* \right\| \leq \omega_F\left(\left(\delta^2 + \delta^{2\mu}\|x_0 - x^*\|_1^2\right)^{1/2} + \delta\right), n \geq N(\delta, F). \tag{2.14}$$

First, let $x_0 \neq x^*$. By (2.11) there exists a sufficiently small number $\delta_0 = \delta_0(F, \xi, \mu) \in \left(0, \widetilde{\delta}(F, \mu)\right) \subset (0,1)$ such that

$$\omega_F\left(\left(\delta^2 + \delta^{2\mu}\|x_0 - x^*\|_1^2\right)^{1/2} + \delta\right) \leq \xi \|x_0 - x^*\|, \forall \delta \in (0, \delta_0). \tag{2.15}$$

If $x_0 = x^*$, then (2.15) should be replaced with $\omega_F\left(\left(\delta^2 + \delta^{2\mu}\|x_0 - x^*\|_1^2\right)^{1/2} + \delta\right) \leq \xi$. Recall that $\|x\| \leq C\|x\|_1, \forall x \in B_1$. Therefore (2.6), (2.7) follow from (2.14) and (2.15). The case $\delta = 0$ can be considered similarly. $\square$

2.2. The finite dimensional case. Consider now the finite dimensional real valued Hilbert space. Compared with subsection 2.1, the main new point here is that a minimizing sequence is replaced with a minimizer, which exists. This case is of our main interest in the current paper because standard piecewise linear finite elements form a finite dimensional space. Unlike the above, we now use the same norm in the regularization term as in the original space. This is because all norms are equivalent in a finite dimensional space. Nevertheless, since the finite dimensional version of the original ill-posed problem "inherits" the ill-posedness, at certain extent, it is still important to use the regularization term for the stabilization.

Let H and H_2 be two real valued Hilbert spaces and $\dim H < \infty$. Norms and scalar products in these spaces denote respectively as $\|\cdot\|, (,), \|\cdot\|_2, (,)_2$. Let $G \subset H$ be an open bounded set and $F: \overline{G} \to H_2$ be a continuous operator. We again consider equations (2.1), (2.2), where $x^* \in G, y, y^* \in H_2$. The functional $M_\alpha(x)$ in (2.3) is now replaced with the functional $J_\alpha(x)$,

$$J_\alpha(x) = \frac{1}{2}\|F(x) - y\|_2^2 + \frac{\alpha}{2}\|x - x_0\|^2, x \in \overline{G}, x_0 \in G. \tag{2.16}$$

The following lemma follows immediately from Weierstrass theorem.

Lemma 2.1. *Let F be the operator defined above in this subsection. Then there exists a regularized solution $x_\alpha \in \overline{G}$,*

$$\inf_{\overline{G}} J_\alpha(x) = \min_{\overline{G}} J_\alpha(x) = J_\alpha(x_\alpha). \tag{2.17}$$

Although a similar result is valid for the case when the set G is unbounded, we do not formulate it here since we do not need it. The following theorem follows immediately from Theorem 2.1.

Theorem 2.2. *Let Hilbert spaces H, H_2, the set $G \subset H$ and the operator $F: \overline{G} \to H_2$ be the same as specified above in this subsection. Also, let (2.2) hods and $\|x_0 - x^*\| \leq 1$. In addition, let the operator F be one-to-one and $\alpha = \alpha(\delta) = \delta^{2\mu}, \mu = const. \in (0, 1/2)$. Let $x_{\alpha(\delta)} \in \overline{G}$ be a regularized solution, i.e. $x_{\alpha(\delta)}$ satisfies (2.17). Let $\xi \in (0,1)$ be an arbitrary constant. Assume first that $x_0 \neq x^*$. Then there exists a sufficiently small number $\delta_0 = \delta_0(F, \mu, \xi) \in (0,1)$ such that for any $\delta \in (0, \delta_0)$*

$$\left\|x_{\alpha(\delta)} - x^*\right\| \leq \xi \|x_0 - x^*\|. \tag{2.18}$$

If $x_0 = x^*$, then (2.18) should be replaced with $\left\|x_{\alpha(\delta)} - x^*\right\| \leq \xi$. In particular, if $\delta = 0$, then δ_0 should be replaced with a sufficiently small number $\alpha_0 \in (0,1)$ and "$\delta \in (0, \delta_0)$" should be replaced with $\alpha \in (0, \alpha_0)$.

3. The Local Strong Convexity of the Functional $J_\alpha(x)$. In this section we prove the local strong convexity of the Tikhonov functional (2.16). Let H and H_2 be two real valued Hilbert spaces. Let scalar products and norms in them be respectively $(,), \|\cdot\|$ and $(,)_2, \|\cdot\|_2$. Let $\mathcal{L}(H, H_2)$ be the the space of all bounded linear operators mapping H into H_2 and let $\|\cdot\|_{\mathcal{L}}$ be the norm in $\mathcal{L}(H, H_2)$. Although we do not assume here that H is finite dimensional, we still use the same norm $\|x - x_0\|$ in the regularization term in (2.16) as the one in the original space H, rather than a stronger norm as in (2.3). This is again because our true goal is to work in a finite dimensional space of finite elements in the adaptivity (section 1). For any $a > 0$ and for any $x \in H$ denote $V_a(x) = \{z \in H : \|x - z\| < a\}$. First, we formulate the following well known theorem.

Theorem 3.1. [38]. *Let $G \subseteq H$ be a convex open set and $L: G \to \mathbb{R}$ be a functional. Suppose that this functional has the Fréchet derivative $L'(x) \in \mathcal{L}(H, \mathbb{R})$ for every point $x \in G$. Then the strong convexity of L on the set G with the strong convexity constant $\rho > 0$ is equivalent with the following condition*

$$(L'(x) - L'(z), x - z) \geq 2\rho \|x - z\|^2, \forall x, z \in G. \tag{3.1}$$

Theorem 3.2. *Let $G \subseteq H$ be a convex open set and $F: \overline{G} \to H_2$ be an operator. Let $x^* \in G$ be an exact solution of equation (2.1) with the exact data y^*. Let $V_1(x^*) \subset G$ and let (2.2) holds. Assume that for* *every $x \in V_1(x^*)$ the operator F has the Fréchet derivative $F'(x) \in \mathcal{L}(H, H_2)$. Suppose that this derivative is uniformly bounded and Lipschitz continuous in $V_1(x^*)$, i.e.*

$$\|F'(x)\|_{\mathcal{L}} \leq N_1,\ \forall x \in V_1(x^*), \tag{3.2}$$

$$\|F'(x) - F'(z)\|_{\mathcal{L}} \leq N_2 \|x - z\|,\ \forall x, z \in V_1(x^*), \tag{3.3}$$

where $N_1, N_2 = const. > 0$. Let

$$\alpha = \alpha(\delta) = \delta^{2\mu}, \quad \forall \delta \in (0, 1), \tag{3.4}$$

$$\mu = const. \in \left(0, \frac{1}{3}\right). \tag{3.5}$$

Then there exists a sufficiently small number $\delta_0 = \delta_0(N_1, N_2, \mu) \in (0, 1)$ such that for all $\delta \in (0, \delta_0)$ the functional $J_\alpha(x)$ with $\alpha = \alpha(\delta)$ in (2.16) is strongly convex in the neighborhood $V_{\delta^{3\mu}}(x^*)$ of x^* with the strong convexity constant $\alpha/4$. In the noiseless case with $\delta = 0$ one should replace $\delta_0 = \delta_0(N_1, N_2, \mu) \in (0, 1)$ with $\alpha_0 = \alpha_0(N_1, N_2) \in (0, 1)$ to be sufficiently small and require that $\alpha \in (0, \alpha_0)$.

Proof. To simplify the presentation, we assume without any loss of generality that in (2.1) $y = 0$. Otherwise, we can replace the operator F with the operator $\widetilde{F}(x) := F(x) - y$. Hence, by (2.2)

$$\|F(x^*)\|_2 \leq \delta. \tag{3.6}$$

For any point $x \in V_1(x^*)$ let $F'^*(x)$ be the linear operator, which is adjoint to the operator $F'(x)$. It can be easily derived from (2.16) that the Fréchet derivative $J'_\alpha(x)$ of the functional $J_\alpha(x)$ acts on the element $u \in H$ as

$$(J'_\alpha(x), u) = (F'^*(x) F(x) + \alpha(x - x_0), u),\ \forall x \in V_1(x^*), \forall u \in H.$$

Hence, $J'_\alpha(x)$ can be considered as an element of the space H,

$$J'_\alpha(x) = F'^*(x) F(x) + \alpha(x - x_0). \tag{3.7}$$

Consider two arbitrary points $x, z \in V_{\delta^{3\mu}}(x^*)$. We have

$$(J'_\alpha(x) - J'_\alpha(z), x - z) = \alpha\|x - z\|^2 + (F'^*(x) F(x) - F'^*(z) F(z), x - z)$$

$$= \alpha\|x - z\|^2 + (F'^*(x) F(x) - F'^*(x) F(z), x - z) + (F'^*(x) F(z) - F'^*(z) F(z), x - z).$$

Hence,

$$(J'_\alpha(x) - J'_\alpha(z), x - z) = \alpha\|x - z\|^2 + A_1 + A_2, \tag{3.8}$$

$$A_1 = (F'^*(x) F(x) - F'^*(x) F(z), x - z),\ A_2 = (F'^*(x) F(z) - F'^*(z) F(z), x - z). \tag{3.9}$$

Estimate numbers A_1, A_2 from the below. We have

$$F(x) - F(z) = \int_0^1 (F'(z + \theta(x - z)) - F'(x))(x - z)\, d\theta. \tag{3.10}$$

Since

$$A_1 = A_1 - (F'^*(x) F'(x)(x - z), x - z) + (F'^*(x) F'(x)(x - z), x - z),$$

then by (**??**) and (3.10)

$$A_1 = \left(F'^* (x) \left(\int_0^1 (F'(z + \theta (x - z)) - F'(x)) (x - z) \, d\theta \right), x - z \right) + (F'^* (x) F'(x) (x - z), x - z). \tag{3.11}$$

Since $\|A\| = \|A^*\|, \forall A \in \mathcal{L}(H, H_2)$, then, using (3.2) and (3.3), we obtain

$$\left| \left(F'^* (x) \left(\int_0^1 (F'(z + \theta (x - z)) - F'(x)) (x - z) \, d\theta \right), x - z \right) \right|$$

$$\leq \|F'(x)\|_{\mathcal{L}} \int_0^1 \|(F'(z + \theta (x - z)) - F'(x)) (x - z)\|_2 \, d\theta \cdot \|x - z\| \leq \frac{1}{2} N_1 N_2 \|x - z\|^3 .$$

Next,

$$(F'^* (x) F'(x) (x - z), x - z) = (F'(x)(x - z), F'(x)(x - z))_2 = \|F'(x)(x - y)\|_2^2 \geq 0.$$

Hence, using (3.11), we obtain

$$A_1 \geq -\frac{1}{2} N_1 N_2 \|x - z\|^3 . \tag{3.12}$$

Now we estimate A_2,

$$|A_2| \leq \|F(z)\|_2 \|F'(x) - F'(z)\|_{\mathcal{L}} \|x - z\| \leq N_2 \|x - z\|^2 \|F(z)\|_2 .$$

By (3.2) and (3.10)

$$\|F(z)\|_2 \leq \|F(z) - F(x^*)\|_2 + \|F(x^*)\|_2 \leq N_1 \|z - x^*\| + \|F(x^*)\|_2 .$$

Hence, using (3.6), we obtain

$$|A_2| \leq N_2 \|x - z\|^2 (N_1 \|z - x^*\| + \|F(x^*)\|_2) \leq N_2 \|x - z\|^2 \left(N_1 \delta^{3\mu} + \delta \right). \tag{3.13}$$

By (3.5) $\delta < \delta^{3\mu}$ for sufficiently small δ. Hence, we can choose $\delta_0 = \delta_0 (N_1, N_2, \mu) \in (0, 1)$ so small that $\left(N_1 \delta^{3\mu} + \delta \right) \leq 2 N_1 \delta^{3\mu}, \forall \delta \in (0, \delta_0)$. Hence, (3.13) implies that

$$A_2 \geq -2 N_2 N_1 \|x - z\|^2 \delta^{3\mu}. \tag{3.14}$$

Hence, using (3.4), (3.8), (3.12) and (3.14) and recalling that $x, z \in V_{\delta^{3\mu}}(x^*)$, we obtain

$$(J'_\alpha (x) - J'_\alpha (z), x - z) \geq \|x - z\|^2 \left[\alpha - \frac{N_1 N_2}{2} \|x - z\| - 2 N_1 N_2 \delta^{3\mu} \right]$$

$$\geq \|x - z\|^2 \left[\delta^{2\mu} - 3 N_1 N_2 \delta^{3\mu} \right] \geq \frac{\delta^{2\mu}}{2} \|x - z\|^2 = \frac{\alpha}{2} \|x - z\|^2 .$$

Combining this with Theorem 3.1, we obtain the assertion of Theorem 3.2. Considerations for the noiseless case are similar. $\square$

Consider now the finite dimensional case.

Theorem 3.3. *Let* $\dim H < \infty, G \subset H$ *be an open bounded convex set, and the rest of conditions of Theorem 3.2 holds with the only exception that (3.5) is now replaced with* $\mu \in (0, 1/4)$. *Let in (2.16) the first guess* x_0 *for the exact solution* x^* *be so accurate that*

$$\|x_0 - x^*\| < \frac{\delta^{3\mu}}{3}. \tag{3.15}$$

Then there exists a sufficiently small number $\delta_0 = \delta_0 (N_1, N_2, \mu) \in (0, 1)$ *such that for every* $\delta \in (0, \delta_0)$ *and for* $\alpha = \alpha(\delta)$ *satisfying (3.4) there exists unique regularized solution* $x_{\alpha(\delta)}$ *of equation (2.1) on the set* G *and* $x_{\alpha(\delta)} \in V_{\delta^{3\mu}}(x^*)$. *In addition, the gradient method of the minimization of the functional* $J_{\alpha(\delta)}(x)$, *which starts at* x_0, *converges to* $x_{\alpha(\delta)}$. *Furthermore, if the operator* F *is one-to-one on* $V_1(x^*)$, *then* $x_{\alpha(\delta)} \in V_{\delta^{3\mu}/3}(x^*)$. *In the noiseless case with* $\delta = 0$ *one should replace* $\delta_0 = \delta_0 (N_1, N_2, \mu) \in (0, 1)$ *with* $\alpha_0 = \alpha_0 (N_1, N_2) \in (0, 1)$ *to be sufficiently small and require that* $\alpha \in (0, \alpha_0)$.

Proof. By Lemma 2.1 there exists a minimizer $x_{\alpha(\delta)} \in \overline{G}$ of the functional $J_{\alpha(\delta)}$. We have $J_{\alpha(\delta)}\left(x_{\alpha(\delta)}\right) \leq J_{\alpha(\delta)}(x^*)$. Also, $\left\|x_{\alpha(\delta)} - x_0\right\| \geq \left\|x_{\alpha(\delta)} - x^*\right\| - \|x_0 - x^*\|$. Hence, using (2.2) and (3.15), we obtain that there exists a sufficiently small number $\delta_0 = \delta_0 (N_1, N_2, \mu) \in (0, 1)$ such that for every $\delta \in (0, \delta_0)$

$$\left\|x_{\alpha(\delta)} - x^*\right\| \leq \frac{\delta}{\sqrt{\alpha}} + 2\|x^* - x_0\| < \delta^{1-\mu} + \frac{2}{3}\delta^{3\mu} = \frac{2}{3}\delta^{3\mu}\left(1 + \frac{3}{2}\delta^{1-4\mu}\right) < \frac{2}{3}\delta^{3\mu} \cdot \frac{3}{2} = \delta^{3\mu}.$$

Hence, $x_{\alpha(\delta)} \in V_{\delta^{3\mu}}(x^*)$. Since by Theorem 3.2 the functional J_α is strongly convex on the set $V_{\delta^{3\mu}}(x^*)$ and the minimizer $x_{\alpha(\delta)} \in V_{\delta^{3\mu}}(x^*)$, then this minimizer is unique. Furthermore, since by (3.15) the point $x_0 \in V_{\delta^{3\mu}}(x^*)$, then it is well known that the gradient method with its starting point at x_0 converges to $x_{\alpha(\delta)}$.

Let now the operator F be one-to-one. Let $\xi \in (0, 1)$ be an arbitrary number and $x_0 \neq x^*$. By Theorem 2.2 we can choose a smaller number $\delta_0 = \delta_0 (N_1, N_2, \mu, \xi)$ such that

$$\left\|x_{\alpha(\delta)} - x^*\right\| \leq \xi\|x_0 - x^*\|, \forall \delta \in (0, \delta_0).$$

Hence, (3.15) implies that $x_{\alpha(\delta)} \in V_{\delta^{3\mu}/3}(x^*)$. If $x_0 = x^*$, then by Theorem 2.2 $\left\|x_{\alpha(\delta)} - x^*\right\| \leq \xi$. Choosing $\xi \in \left(0, \delta^{3\mu}/3\right)$, we again obtain that $x_{\alpha(\delta)} \in V_{\delta^{3\mu}/3}(x^*)$. The noiseless case is similar. $\square$

4. The Space of Finite Elements. To prove the relaxation property of the adaptivity, we need to introduce the space of finite elements. We consider only standard piecewise linear finite elements, which are triangles in 2-d and tetrahedra in 3-d. Let $\Omega \subset \mathbb{R}^n, n = 2, 3$ be a bounded domain. Consider a triangulation T of Ω with a rather coarse mesh. We obtain a polygonal domain D and assume for brevity that $D = \Omega$. Following section 76.4 of [22], we associate with triangulation T piecewise linear functions $\{e_j(x, T)\}_{j=1}^{p(T)} \subset C\left(\overline{\Omega}\right)$, which are called global *test functions*. Functions $\{e_j(x, T)\}_{j=1}^{p(T)}$ are linearly independent in Ω. The number $p(T)$ of these global functions equals to the number of the mesh points in the domain Ω.

Let $\{N_i\} = N_1, N_2, ..., N_{p(T)}$ be the enumeration for nodes in the triangulation T. Then test functions should satisfy to the following condition for all $i, j \in \{N_i\}$

$$e_j(N_i, T) = \begin{cases} 1, i = j, \\ 0, i \neq j. \end{cases}$$

The linear space of finite elements with its basis $\{e_j(x, T)\}_{j=1}^{p(T)}$ is defined as

$$V(T) = \left\{v(x) : v \in C\left(\overline{\Omega}\right), v\mid_K \text{ is linear on } K\ \forall K \in T\right\},$$

where $v\mid_K$ is the restriction of the function v on the element K. Each function $v \in V(T)$ can be represented as

$$v(x) = \sum_{j=1}^{p(T)} v(N_j)e_j(x, T).$$

Let $h(K_j)$ be the diameter of the triangle/tetrahedra $K_j \subset T$. Then the number h,

$$h = \max_{K_j \subset T} h(K_j)$$

is called the *maximal grid step size* of the triangulation T. Let r be the radius of the maximal circle/sphere inscribed in K_j. We impose the shape regularity assumption for all triangles/tetrahedra uniformly for all possible triangulations T which we consider. Specifically, we assume that

$$a_1 \leqslant h(K_j) \leqslant r a_2, \quad a_1, a_2 = const. > 0, \ \forall K_j \subset T, \ \forall \ T, \tag{4.1}$$

where numbers a_1, a_2 are independent on the triangulation T. Obviously, the number of all possible triangulations satisfying (4.1) is finite. Thus, we introduce the following finite dimensional linear space H,

$$H = \bigcup_T Span(V(T)), \ \forall T \ \text{ satisfying (4.1).}$$

Hence,

$$\dim H < \infty, \ H \subset \left(C\left(\overline{\Omega}\right) \cap H^1(\Omega) \right), \ \partial_{x_i} f \in L_\infty(\Omega), \ \forall f \in H. \tag{4.2}$$

In (4.2) "$\subset$" means the inclusion of sets. We equip H with the same inner product as the one in $L_2(\Omega)$. Denote $(,)$ and $\|\cdot\|$ the inner product and the norm in H respectively, $\|f\|_H := \|f\|_{L_2(\Omega)} := \|f\|, \forall f \in H$.

Keeping in mind the mesh refinement process in the adaptivity, we now explain how do we construct triangulations $\{T_n\}$ as well as corresponding subspaces $\{M_n\}$ of the space H which correspond to mesh refinements. Consider the first triangulation T_1 with rather coarse mesh. We set $M_1 := V(T_1) \subset H$. Suppose that the pair (T_n, M_n) is constructed after n mesh refinements and that the basis functions in the space M_n are $\{e_j(x, T_n)\}_{j=1}^{p(T_n)}$. We now want to refine the mesh again. We define the pair (T_{n+1}, M_{n+1}) as follows. First, we refine the mesh in the standard manner as it is usually done when working with triangular/tetrahedron finite elements. When doing so, we keep (4.1). Hence, we obtain both the triangulation T_{n+1} and the corresponding test functions $\{e_j(x, T_{n+1})\}_{j=1}^{p(T_{n+1})}$. It is well known that test functions $\{e_j(x, T_n)\}_{j=1}^{p(T_n)}$ depend linearly on new test functions $\{e_j(x, T_{n+1})\}_{j=1}^{p(T_{n+1})}$. Thus, we define the subspace M_{n+1} as

$$M_{n+1} := Span\left(\{e_j(x, T_{n+1})\}_{j=1}^{p(T_{n+1})} \right).$$

Therefore, we have obtained a finite set of linear subspaces $\{M_n\}_{n=1}^N$ of the space H. Each subspace M_n corresponds to the mesh refinement number n and

$$M_n \subset M_{n+1} \subset H, n \in [1, N].$$

Let I be the identity operator on H. For any subspace $M \subset H$, let $P_M : H \to M$ be the orthogonal projection operator onto M. Denote for brevity $P_n := P_{M_n}$. Let h_n be the mesh function for T_n defined as a maximal diameter of the elements in triangulation T_n. Let f_n^I be the standard interpolant of the function $f \in H$ on triangles/tetrahedra of T_n, see section 76.4 of [22]. It can be easily derived from formula (76.3) of [22] that

$$\left\| f - f_n^I \right\|_{L_\infty(\Omega)} \leq K \left\| \nabla f \right\|_{L_\infty(\Omega)} h_n, \forall f \in H, \tag{4.3}$$

where $K = K(\Omega, r, a_1, a_2) = const. > 0$. Since $f_n^I \in H, \forall f \in H$, then by one of well known properties of orthogonal projection operators,

$$\| f - P_n f \| \leqslant \left\| f - f_n^I \right\|, \ \forall f \in H. \tag{4.4}$$

Hence, (4.3) and (4.4) imply that

$$\|f - P_n f\|_{L_\infty(\Omega)} \leqslant K \|\nabla f\|_{L_\infty(\Omega)} h_n, \forall f \in H. \tag{4.5}$$

Since H is a finite dimensional space in which all norms are equivalent, it is convenient for us to rewrite (4.5) with a different constant K in the form

$$\|x - P_n x\| \leqslant K \|x\| h_n, \ \forall x \in H. \tag{4.6}$$

Remark 4.1. Everywhere below H is the space defined in this section. We view the space H as an "ideal" space of very fine finite elements, which cannot be reached in practical computations. At the same time, all other spaces of finite elements we work with below are subspaces of H. In particular, this means that we assume without further mentioning that (4.4) is valid for all grid step sizes considered below. We also assume that we refine meshes not only until condition (4.4) is fulfilled, but also check the condition

$$h(K_{min})/h(K_{max}) \leq C \tag{4.7}$$

with the some constant C, where $h(K_{min}), h(K_{max})$ are the smallest and the largest diameters for the elements in the mesh T_n, correspondingly. We perform local mesh refinements automatically and check the condition (4.7) numerically.

5. Relaxation. Since we sequentially minimize the Tikhonov functional on subspaces $\{M_n\}_{n=1}^{N}$ in the adaptivity procedure, then we need to establish first the existence of a minimizer on each of these subspaces. In this section the set G and the operator F are the same as in Theorem 3.3, and the functional $J_\alpha(x)$ is the same as in (2.16). Theorem 5.1 ensures both existence and uniqueness of the minimizer of the functional J_α on each subspace of the space H, as long as the maximal grid step size of finite elements, which are involved in that subspace, is sufficiently small, although condition (4.1) is still satisfied, see Remark 4.1. The smallness of that grid step size depends on the upper estimate of the norm $\|x^*\|$ of the exact solution. Since by the fundamental concept of Tikhonov (section 1.4 of [11]), one can assume an *a priori* knowledge of this estimate, then this assumption imposes an upper bound on the maximal grid step size.

Theorem 5.1. *Let $G \subset H$ be an open bounded convex set, $F : \overline{G} \to H_2$ be an operator and J_α be the functional defined in (2.16), where H_2 is a real valued Hilbert space. Let $x^* \in G$ be the exact solution of equation (2.2) with the exact data y^* and $\delta \in (0,1)$ be the level of the error in the data, as in (2.2). Suppose that $V_1(x^*) \subset G$. Assume that for every $x \in V_1(x^*)$ the operator F has the Frechét derivative $F'(x)$ and that this derivative is uniformly bounded and Lipschitz continuous in $V_1(x^*)$, i.e.*

$$\|F'(x)\|_{\mathcal{L}} \leq N_1, \ \forall x \in V_1(x^*), \tag{5.1}$$

$$\|F'(x) - F'(z)\|_{\mathcal{L}} \leq N_2 \|x - z\|, \ \forall x, z \in V_1(x^*), \tag{5.2}$$

where $N_1, N_2 = const. \geq 1$. Let

$$\alpha = \alpha(\delta) = \delta^{2\mu}, \quad \forall \delta \in (0,1), \tag{5.3}$$

$$\mu = const. \in \left(0, \frac{1}{4}\right). \tag{5.4}$$

Let $M \subseteq H$ be a subspace of H and let $V_{\delta^{3\mu}}(x^*) \cap M \neq \varnothing$. Assume that $\|x^*\| \leq A$, where the number A is known in advance. Suppose that the maximal grid step size $\widetilde{h}$ of finite elements of M be so small that (also, see Remark 4.1)

$$\widetilde{h} \leq \frac{\delta^{4\mu}}{5AN_2K}, \tag{5.5}$$

where K is the constant in (4.6). Furthermore, assume that the first guess x_0 for the exact solution x^* in the functional $J_{\alpha(\delta)}$ is so accurate that

$$\|x_0 - x^*\| < \frac{\delta^{3\mu}}{3}.$$

Then there exists a sufficiently small number $\delta_0 = \delta_0(N_1, N_2, \mu) \in (0,1)$ such that for every $\delta \in (0, \delta_0)$ there exists unique minimizer $x_{M,\alpha(\delta)} \in G \cap M$ of the functional J_α on the set $G \cap M$ and $x_{M,\alpha(\delta)} \in V_{\delta^{3\mu}}(x^*) \cap M$. Furthermore, the functional $J_\alpha(x)$ is strongly convex on the set $V_{\delta^{3\mu}}(x^*) \cap M$ with the strong convexity constant $\alpha(\delta)/4$. In particular, let $x_{\alpha(\delta)} \in V_{\delta^{3\mu}/3}(x^*)$ be the regularized solution of equation (2.1) (Theorem 3.3). Then the following a posteriori error estimate holds

$$\left\| x_{M,\alpha(\delta)} - x_{\alpha(\delta)} \right\| \leq \frac{2}{\delta^{2\mu}} \left\| J'_\alpha\left(x_{M,\alpha(\delta)}\right)\right\|.$$

Note that since in Theorem 5.1 $V_1(x^*) \subset G$ and $V_{\delta^{3\mu}}(x^*) \cap M \neq \varnothing$, then $G \cap M \neq \varnothing$. We do not prove this theorem here and refer instead to Theorem 4.9.2 of [11]; also see Theorem 3.2 of [15] for a similar result.

Theorem 5.2 (relaxation property of the adaptivity)**.** *Let $M_n \subset H$ be the subspace obtained after n mesh refinements, as described in section 4. Let h_n be the maximal grid step size of the subspace M_n. Suppose that all conditions of Theorem 5.1 hold with the only exception that the subspace M is replaced with M_n and the inequality (5.5) is replaced with (also, see Remark 4.1)*

$$h_n \leq \frac{\delta^{4\mu}}{5AN_2K}. \tag{5.6}$$

Also, let $V_{\delta^{3\mu}}(x^*) \cap M_1 \neq \varnothing$. Let $x_n \in V_{\delta^{3\mu}}(x^*) \cap M_n$ be the unique minimizer of the functional $J_\alpha(x)$ in (2.16) on the set $G \cap M_n$ (Theorem 5.1) and $x_{\alpha(\delta)} \in V_{\delta^{3\mu}/3}(x^*)$ be the unique regularized solution (Theorem 3.3). Assume that

$$x_n \neq x_{\alpha(\delta)}, \tag{5.7}$$

i.e. $x_{\alpha(\delta)} \notin M_n$, meaning that the regularized solution is not yet reached after n mesh refinements. Let $\eta \in (0,1)$ be an arbitrary number. Then one can choose the maximal grid size $h_{n+1} = h_{n+1}(N_1, N_2, \delta, A, K, \eta) \in (0, h_n]$ of the mesh refinement number $(n+1)$ so small that

$$\left\| x_{n+1} - x_{\alpha(\delta)} \right\| \leq \eta \left\| x_n - x_{\alpha(\delta)} \right\|, \tag{5.8}$$

where $x_{n+1} \in V_{\delta^{3\mu}}(x^*) \cap M_{n+1}$ is the unique minimizer of the functional (2.16) on the set $G \cap M_{n+1}$ (also, see Remark 4.1). Hence,

$$\left\| x_{n+1} - x_{\alpha(\delta)} \right\| \leq \eta^n \left\| x_1 - x_{\alpha(\delta)} \right\|. \tag{5.9}$$

Proof. Since $V_{\delta^{3\mu}}(x^*) \cap M_1 \neq \varnothing, M_1 \subseteq M_n$ and $V_{\delta^{3\mu}}(x^*) \subset V_1(x^*) \subset G$, then $(V_{\delta^{3\mu}}(x^*) \cap M_n) \subset (V_{\delta^{3\mu}}(x^*) \cap M_{n+1}) \neq \varnothing$ and also $(V_1(x^*) \cap M_n) \subset (V_1(x^*) \cap M_{n+1}) \neq \varnothing$. In particular, $(G \cap M_n) \subset (G \cap M_{n+1}) \neq \varnothing$. Next, by Theorem 5.1 the functional (2.16) is strongly convex on the set $V_{\delta^{3\mu}}(x^*) \cap M_{n+1}$ with the strong convexity constant $\alpha/4$. Hence, Theorem 3.1 implies that

$$\frac{\alpha}{2}\left\| x_{n+1} - x_{\alpha(\delta)} \right\|^2 \leq \left(J'_\alpha(x_{n+1}) - J'_\alpha\left(x_{\alpha(\delta)}\right), x_{n+1} - x_{\alpha(\delta)} \right). \tag{5.10}$$

Since x_{n+1} is the minimizer on $G \cap M_{n+1}$ and x_α is the minimizer on the set G, then

$$\left(J'_\alpha(x_{n+1}), z \right) = 0, \ \forall z \in M_{n+1}; J'_\alpha\left(x_{\alpha(\delta)}\right) = 0. \tag{5.11}$$

Relations (5.11) justify the application of the Galerkin orthogonality principle, see, e.g. [4, 6]. By (5.11)

$$\left(J'_\alpha(x_{n+1}) - J'_\alpha\left(x_{\alpha(\delta)}\right), x_{n+1} - P_{n+1}x_{\alpha(\delta)} \right) = 0. \tag{5.12}$$

Next,

$$x_{n+1} - x_{\alpha(\delta)} = \left(x_{n+1} - P_{n+1}x_{\alpha(\delta)} \right) + \left(P_{n+1}x_{\alpha(\delta)} - x_{\alpha(\delta)} \right).$$

Hence, (5.10) and (5.12) imply that

$$\frac{\alpha}{2}\left\|x_{n+1}-x_{\alpha(\delta)}\right\|^{2}\leq\left(J_{\alpha}^{\prime}\left(x_{n+1}\right)-J_{\alpha}^{\prime}\left(x_{\alpha(\delta)}\right),P_{n+1}x_{\alpha(\delta)}-x_{\alpha(\delta)}\right). \tag{5.13}$$

It follows from (3.7) that conditions (5.1) and (5.2) imply that

$$\left\|J_{\alpha}^{\prime}\left(x_{n+1}\right)-J_{\alpha}^{\prime}\left(x_{\alpha(\delta)}\right)\right\|\leq N_{3}\left\|x_{n+1}-x_{\alpha(\delta)}\right\| \tag{5.14}$$

with a constant $N_3=N_3\left(N_1,N_2\right)>0$. Also, by (4.6)

$$\left\|x_{\alpha(\delta)}-P_{n+1}x_{\alpha(\delta)}\right\|\leq K\left\|x_{\alpha(\delta)}\right\|h_{n+1}. \tag{5.15}$$

Using the Cauchy-Schwarz inequality as well as (5.3), (5.14) and (5.15), we obtain from (5.13)

$$\left\|x_{n+1}-x_{\alpha(\delta)}\right\|\leq\frac{2KN_{3}}{\delta^{2\mu}}\left\|x_{\alpha(\delta)}\right\|h_{n+1}. \tag{5.16}$$

Since by one of conditions of Theorem 5.1 we have an *a priori* known upper estimate $\|x^*\|\leq A$, we now can estimate the norm $\|x_\alpha\|$. Since by Theorem 3.3 $x_{\alpha(\delta)}\in V_{\delta^{3\mu}/3}\left(x^*\right)$, then

$$\left\|x_{\alpha(\delta)}\right\|\leq\left\|x_{\alpha(\delta)}-x^{*}\right\|+\left\|x^{*}\right\|\leq\frac{\delta^{3\mu}}{3}+A.$$

Hence, (5.16) becomes

$$\left\|x_{n+1}-x_{\alpha(\delta)}\right\|\leq\frac{2KN_{3}}{\delta^{2\mu}}\left(\frac{\delta^{3\mu}}{3}+A\right)h_{n+1}. \tag{5.17}$$

Let $\eta\in(0,1)$ be an arbitrary number. Since $\left\|x_n-x_{\alpha(\delta)}\right\|\neq0$, then we can choose $h_{n+1}=h_{n+1}\left(N_2,\delta,A,K\right)\in\left(0,h_n\right]$ so small that

$$\frac{2KN_{3}}{\delta^{2\mu}}\left(\frac{\delta^{3\mu}}{3}+A\right)h_{n+1}\leq\eta\left\|x_{n}-x_{\alpha(\delta)}\right\|. \tag{5.18}$$

Comparing (5.18) with (5.17), we obtain the target estimate (5.8). $\square$

Theorem 5.3 follows immediately from Theorem 2.2 and (5.9).

Theorem 5.3. *Let all conditions of Theorem 5.2 hold. Let $\xi\in(0,1)$ be an arbitrary number. Then there exists a sufficiently small number $\delta_0=\delta_0\left(N_1,N_2,\mu,\xi\right)\in(0,1)$ and a decreasing sequence of maximal grid step sizes $\{h_k\}_{k=1}^{n+1}$ such that*

$$\left\|x_{k+1}-x^{*}\right\|\leq\eta^{k}\left\|x_{1}-x_{\alpha(\delta)}\right\|+\xi\left\|x_{0}-x^{*}\right\|,k=1,...,n. \tag{5.19}$$

Since h_{n+1} is the maximal grid step size in the entire domain Ω, it seems to be at the first glance that relaxation Theorems 5.2 and 5.3 are about mesh refinements in the entire domain Ω rather than about local mesh refinements in subdomains, as it is the case in the adaptivity. Assuming that conditions of Theorem 5.3 hold, we now show that local mesh refinements are also covered by this theorem. Suppose that the first guess x_0 is sufficiently close to x^*, as it should be in the realistic case, because of the problem of local minima and ravines of the Tikhonov functional, see section 1. Hence, we need to refine x_0 via the adaptivity technique. Suppose that the domain Ω is split in two subdomains, $\Omega=\Omega_1\cup\Omega_2,\Omega_1\cap\Omega_2=\varnothing$. Assume that the function x_0 is changing slowly in Ω_1 and has some "bumps" in Ω_2. These bumps correspond to small inclusions. It is these inclusions rather than slowly changing functions, which are of the main applied interest in imaging. Indeed, those small abnormalities model, e.g. land mines, tumors, etc. Hence, it is reasonable to assume that x^* is also changing slowly in Ω_1. Next, because of Theorem 2.2 and because all norms in H

are equivalent, it is also reasonable to assume that the regularized solution x_α is changing slowly in Ω_1. As the *simplest* example, one can assume that

$$\nabla x_{\alpha(\delta)} = \nabla x^* = 0 \text{ in } \Omega_1. \tag{5.20}$$

A more general case can be considered along the same lines. Thus, inequality (5.21) of Theorem 5.4 is a reasonable one. Furthermore, it is reasonable to assume that mesh refinements do not take place in Ω_1, but only in Ω_2. Let $h^{(1)}$ be the maximal grid step size in Ω_1.

Theorem 5.4 (relaxation of the adaptivity for local mesh refinements). *Assume that conditions of Theorem 5.2 hold. Then there exists a sufficiently small number* $\delta_0 = \delta_0(N_1, N_2, \mu, \xi) \in (0,1)$ *and a decreasing sequence of maximal grid step sizes* $\left\{\widetilde{h}_k\right\}_{k=1}^{n+1}$ *such that if* $\left\|\nabla x_{\alpha(\delta)}\right\|_{L_\infty(\Omega_1)}$ *is so small that*

$$\frac{2KN_3}{\delta^{2\mu}} \left\|\nabla x_{\alpha(\delta)}\right\|_{L_\infty(\Omega_1)} h^{(1)} \leq \frac{\eta}{2} \left\|x_k - x_{\alpha(\delta)}\right\|, k = 1, ..., n, \tag{5.21}$$

then (5.19) holds with the replacement of $\{h_k\}_{k=1}^{n+1}$ *with* $\left\{\widetilde{h}_k\right\}_{k=1}^{n+1}$.

Proof. By (5.13) and (5.14)

$$\begin{aligned} \left\|x_{k+1} - x_{\alpha(\delta)}\right\| \leq \frac{2N_3}{\delta^{2\mu}} \left\|x_{\alpha(\delta)} - P_{k+1}x_{\alpha(\delta)}\right\| = \\ \frac{2N_3}{\delta^{2\mu}} \left(\left\|x_{\alpha(\delta)} - P_{k+1}x_{\alpha(\delta)}\right\|_{L_2(\Omega_1)} + \left\|x_{\alpha(\delta)} - P_{k+1}x_{\alpha(\delta)}\right\|_{L_2(\Omega_2)}\right). \end{aligned} \tag{5.22}$$

By (4.5) and (5.21)

$$\frac{2N_3}{\delta^{2\mu}} \left\|x_{\alpha(\delta)} - P_{k+1}x_{\alpha(\delta)}\right\|_{L_2(\Omega_1)} \leq \frac{2KN_3}{\delta^{2\mu}} \left\|\nabla x_{\alpha(\delta)}\right\|_{L_\infty(\Omega_1)} h^{(1)} \leq \frac{\eta}{2} \left\|x_k - x_{\alpha(\delta)}\right\|. \tag{5.23}$$

Next, we obtain similarly with (5.18)

$$\frac{2N_3}{\delta^{2\mu}} \left\|x_{\alpha(\delta)} - P_{k+1}x_{\alpha(\delta)}\right\|_{L_2(\Omega_2)} \leq \frac{\eta}{2} \left\|x_k - x_{\alpha(\delta)}\right\|. \tag{5.24}$$

It follows from (5.22)-(5.24) that

$$\left\|x_{k+1} - x_{\alpha(\delta)}\right\| \leq \eta \left\|x_k - x_{\alpha(\delta)}\right\|, k = 1, ..., n.$$

Hence, (5.9) holds. Finally, (5.19) follows from (5.9) and Theorem 2.2. □

Remark 5.1. Thus, Theorem 5.3 claims that the accuracy of the reconstruction of the exact solution x^* improves with mesh refinements, i.e. the relaxation takes place. Comparison of (2.18) with (5.7) and (5.19) shows that the solution accuracy improvement continues until reaching the regularized solution $x_{\alpha(\delta)}$. It is important that the accuracy of $x_{\alpha(\delta)}$ is better than the accuracy of the first guess x_0. Indeed, this ensures that it is worthy to undertake the "adaptivity effort" to improve the accuracy of the regularized solution. Theorem 5.4 implies that the same is true for local mesh refinements.

6. Adaptivity for a Coefficient Inverse Problem. We now reformulate some of above theorems for the case of a specific CIP. To save space, we do not prove theorems of this section. Instead, we point to those results of Chapter 4 of [11] from which these theorems follow easily.

6.1. Coefficient Inverse Problem and Tikhonov functional. Let $\Omega \subset \mathbb{R}^3$ be a convex bounded domain with the boundary $\partial\Omega \in C^3$. Let the point $x_0 \notin \overline{\Omega}$. For $T > 0$ denote $Q_T = \Omega \times (0,T), S_T = \partial\Omega \times (0,T)$. Let $d > 1$ be a certain number, $\omega \in (0,1)$ be a sufficiently small number, and the function $c(x) \in C(\mathbb{R}^3)$ be such that

$$c(x) \in (1 - \omega, d + \omega) \text{ in } \Omega, c(x) = 1 \text{ outside of } \Omega. \tag{6.1}$$

Below we specify $c(x)$ more. Consider the solution $u(x,t)$ of the following Cauchy problem

$$c(x)\, u_{tt} = \Delta u, x \in \mathbb{R}^3, t \in (0,T), \tag{6.2}$$

$$u(x,0) = 0, u_t(x,0) = \delta(x - x_0). \tag{6.3}$$

Equation (6.2) governs propagation of acoustic waves, in which case $c(x) = 1/b^2(x)$, where $b(x)$ is the sound speed and $u(x,t)$ is the amplitude of the acoustic wave [44]. In addition, (6.2) governs propagation of the electromagnetic field in 2-d, in which case $c(x) = \varepsilon_r(x)$ is the spatially distributed dielectric constant and $u(x,t)$ is one of components of the electric field [41]. Although in the latter application equation (6.2) is valid only in 2-d, we have successfully used this equation to model experimental data, which are obviously in 3-d, see [11, 14, 31] and section 8.

Remark 6.1. An alternative to the point source in (6.3) is the incident plane wave in the case when it is initialized at the plane $\{x_3 = x_{3,0}\}$ such that $\{x_3 = x_{3,0}\} \cap \overline{\Omega} = \varnothing$. The formalism of derivations below is similar in this case. In our derivations below we focus on (6.3), because this is the most convenient case for derivations. However, in numerical studies we use the incident plane wave, because this case has shown a better performance than the case (6.3) of the point source.

Coefficient Inverse Problem (CIP**)**. Let conditions (6.1)-(6.3) hold. Assume that the coefficient $c(x)$ is unknown inside the domain Ω. Determine this coefficient for $x \in \Omega$, assuming that the following function $g(x,t)$ is known

$$u\mid_{S_T} = g(x,t). \tag{6.4}$$

The function $g(x,t)$ can be interpreted as the result of measurements of the outcome wave field $u(x,t)$ at the boundary of the domain of interest Ω. Since the function $c(x) = 1$ outside of Ω, then (6.2)-(6.4) imply

$$u_{tt} = \Delta u, (x,t) \in \left(\mathbb{R}^3 \diagdown \Omega\right) \times (0,T),$$
$$u(x,0) = u_t(x,0) = 0, x \in \mathbb{R}^3 \diagdown \Omega, u\mid_{S_T} = g(x,t).$$

Solving this initial boundary value problem for $(x,t) \in \left(\mathbb{R}^3 \diagdown \Omega\right) \times (0,T)$, we uniquely and stably obtain Neumann boundary condition $p(x,t)$ for the function u,

$$\partial_n u\mid_{S_T} = p(x,t). \tag{6.5}$$

CIPs are quite complex problems. Hence, to handle them, one naturally needs to impose some simplifying assumptions. In this particular CIP our theory is not working unless we replace the $\delta-$function in (6.3) by a smooth function, which approximates $\delta(x - x_0)$ in the distribution sense. Let $\varkappa \in (0,1)$ be a sufficiently small number. We replace $\delta(x - x_0)$ in (6.3) with the function $\delta_\varkappa(x - x_0)$,

$$\delta_\varkappa(x - x_0) = \begin{cases} C_\varkappa \exp\left(\frac{1}{|x-x_0|^2 - \varkappa^2}\right), |x - x_0| < \varkappa, \\ 0, |x - x_0| > \varkappa, \end{cases} \quad \int\limits_{\mathbb{R}^3} \delta_\varkappa(x - x_0)\, dx = 1. \tag{6.6}$$

We assume that $\varkappa$ is so small that

$$\delta_\varkappa(x - x_0) = 0 \text{ in } \overline{\Omega}. \tag{6.7}$$

Let $\zeta \in (0,1)$ be a sufficiently small number. Consider the function $z_\zeta \in C^\infty[0,T]$ such that

$$z_\zeta(t) = \begin{cases} 1, t \in [0, T - 2\zeta], \\ 0, t \in [T - \zeta, T], \\ \text{between 0 and 1 for } t \in [0, T - 2\zeta, T - \zeta]. \end{cases} \tag{6.8}$$

We now introduce state and adjoint problems.

State Problem. Find the solution $v(x,t)$ of the following initial boundary value problem

$$\begin{aligned} c(x)\,v_{tt} - \Delta v &= 0 \text{ in } Q_T, \\ v(x,0) = v_t(x,0) &= 0, \\ \partial_n v \mid_{S_T} &= p(x,t). \end{aligned} \tag{6.9}$$

Adjoint Problem. Find the solution $\lambda(x,t)$ of the following initial boundary value problem with the reversed time

$$\begin{aligned} c(x)\,\lambda_{tt} - \Delta \lambda &= 0 \text{ in } Q_T, \\ \lambda(x,T) = \lambda_t(x,T) &= 0, \\ \partial_n \lambda \mid_{S_T} &= z_\zeta(t)\,(g-v)(x,t). \end{aligned} \tag{6.10}$$

Here functions $v \in H^1(Q_T)$ and $\lambda \in H^1(Q_T)$ are weak solutions of problems (6.9) and (6.10) respectively. In fact, we need a higher smoothness of these functions, which we specify below. In (6.10) and (6.9) functions g and p are the ones from (6.4) and (6.5) respectively. Hence, to solve the adjoint problem, one should solve the state problem first. The function $z_\zeta(t)$ is introduced to ensure the validity of compatibility conditions at $\{t=T\}$ in (6.10). The Tikhonov functional for the above CIP is

$$E_\alpha(c) = \frac{1}{2}\int\limits_{S_T} (v\mid_{S_T} - g(x,t))^2 z_\zeta(t)\,d\sigma dt + \frac{1}{2}\alpha \int\limits_{\Omega} (c - c_{glob})^2 dx, \tag{6.11}$$

where c_{glob} is the approximate solution obtained by our approximately globally convergent numerical method on the first stage of our two stage numerical procedure (section 1) and α is the small regularization parameter as above. To figure out the Fréchet derivative of the functional $E_\alpha(c)$, we introduce the Lagrangian $L(c)$,

$$L(c) = E_\alpha(c) - \int\limits_{Q_T} c(x) v_t \lambda_t dx dt + \int\limits_{Q_T} \nabla v \nabla \lambda dx dt - \int\limits_{S_T} p\lambda d\sigma_x dt. \tag{6.12}$$

The sum of integral terms in $L(c)$ equals zero, because of the definition of the weak solution $v \in H^1$ of the problem (6.9) as $v(x,0)=0$ and

$$\int\limits_{Q_T} (-c(x) v_t w_t + \nabla v \nabla w)\,dx dt - \int\limits_{S_T} p w d\sigma_x dt = 0, \forall w \in H^1(Q_T), w(x,T) = 0, \tag{6.13}$$

see section 5 of Chapter 4 of [36]. Hence, $L(c) = E_\alpha(c)$. It is more convenient to calculate the Fréchet derivative of the Tikhonov functional $E_\alpha(c)$ written in the form (6.12) rather than in the form (6.11). Because of this, it is necessary to consider Fréchet derivatives of functions v, λ with respect to the coefficient c (in certain functional spaces). This in turn requires to establish a higher smoothness of functions v, λ than just $H^1(Q_T)$ [11, 13].

State and adjoint problems are concerned only with the domain Ω rather than with the entire space $\mathbb{R}^3$. We define the space Z as

$$Z = \left\{ f : f \in C\left(\overline{\Omega}\right) \cap H^1(\Omega), c_{x_i} \in L_\infty(\Omega), i = 1,2,3 \right\}, \|f\|_Z = \|f\|_{C(\overline{\Omega})} + \sum_{i=1}^{3} \|f_{x_i}\|_{L_\infty(\Omega)}.$$

Clearly $H \subset Z$ as a set. To apply the theory of above sections, we express in subsection 6.2 the function $c(x)$ via standard piecewise linear finite elements. Hence, we assume below that $c \in Y$, where

$$Y = \{c \in Z : c \in (1-\omega, d+\omega)\}. \tag{6.14}$$

Theorem 6.1 can be easily derived from a combination of Theorems 4.7.1, 4.7.2 and 4.8 of [11] as well as from Theorems 3.1, 3.2 of [13].

Theorem 6.1. *Let $\Omega \subset \mathbb{R}^3$ be a convex bounded domain with the boundary $\partial\Omega \in C^2$ and such that there exists a function $a \in C^2\left(\overline{\Omega}\right)$ such that $a \mid_{\partial\Omega}= 0, \partial_n a \mid_{\partial\Omega}= 1$. Assume that there exists functions $P(x,t), \Phi(x,t)$ such that*

$$P \in H^6(Q_T), \Phi \in H^5(Q_T); \partial_n P \mid_{S_T}= p(x,t), \partial_n \Phi \mid_{S_T}= z_\zeta(t)\, g(x,t),$$
$$\partial_t^j P(x,0) = \partial_t^j \Phi(x,0) = 0, j = 1,2,3,4.$$

Then for every function $c \in Y$ functions $v, \lambda \in H^2(Q_T)$, where v, λ are solutions of state and adjoint problems (6.9), (6.10). Also, for every $c \in Y$ there exists Fréchet derivative $E'_\alpha(c)$ of the Tikhonov functional $E_\alpha(c)$ in (6.11) and

$$E'_\alpha(c)(x) = \alpha(c - c_{glob})(x) - \int_0^T (u_t\lambda_t)(x,t)\ dt := \alpha(c - c_{glob})(x) + y(x). \tag{6.15}$$

The function $y(x) \in C\left(\overline{\Omega}\right)$ and the there exists a constant $B = B(\Omega, a, d, \omega, z_\zeta) > 0$ such that

$$\|y\|_{C(\overline{\Omega})} \leq \|c\|^2_{C(\overline{\Omega})} \exp(BT)\left(\|P\|^2_{H^6(Q_T)} + \|\Phi\|^2_{H^5(Q_T)}\right). \tag{6.16}$$

The functional of the Fréchet derivative $E'_\alpha(c)$ acts on any function $b \in Z$ as

$$E'_\alpha(c)(b) = \int_\Omega E'_\alpha(c)(x)\, b(x)\, dx.$$

6.2. Relaxation property of mesh refinements for the functional $E_\alpha(c)$. We now need to "project" the relaxation property for the specific functional $E_\alpha(c)$ for our CIP. To do this, we need to define first the operator F for our specific case. Let Y be the set of functions defined in (6.14) and H be the finite dimensional space of finite elements constructed in section 4 (also, see Remark 4.1). We define the set G as $G := Y \cap H$. We consider the set G as the subset of the space H with the same norm as in H. In particular, $\overline{G} = \left\{c(x) \in H : c(x) \in [1-\omega, d+\omega] \text{ for } x \in \overline{\Omega}\right\}$. Let the Hilbert space $H_2 := L_2(S_T)$. We define the operator F as

$$F : \overline{G} \to H_2, F(c)(x,t) = z_\zeta(t)\left[g(x,t) - v(x,t,c)\right],\ \ (x,t) \in S_T, \tag{6.17}$$

where the function $v := v(x,t,c)$ is the weak solution (6.13) of the state problem (6.9), g is the function in (6.4) and $z_\zeta(t)$ is the function defined in (6.8). For any function $b \in H$ consider the weak solution $\widetilde{u}(x,t,c,b) \in H^1(Q_T)$ of the following initial boundary value problem

$$c(x)\,\widetilde{u}_{tt} = \Delta\widetilde{u} - b(x)\, v_{tt}, (x,t) \in Q_T,$$
$$\widetilde{u}(x,0) = \widetilde{u}_t(x,0) = 0, \widetilde{u} \mid_{S_T}= 0.$$

Theorem 6.2 can be easily derived from a combination of Theorems 4.7.2 and 4.10 of [11].

Theorem 6.2. *Let $\Omega \subset \mathbb{R}^3$ be a convex bounded domain with the boundary $\partial\Omega \in C^2$. Suppose that there exist functions $a(x), P(x,t), \Phi(x,t)$ satisfying conditions of Theorem 6.1. Then the function $\widetilde{u}(x,t,c,b) \in H^2(Q_T), \forall c \in \overline{G}, \forall b \in H$. Also, the operator F in (6.17) has the Fréchet derivative $F'(c)(b)$,*

$$F'(c)(b) = -z_\zeta(t)\,\widetilde{u}(x,t,c,b) \mid_{S_T}, \forall c \in \overline{G}, \forall b \in H.$$

Let $B = B(\Omega, a, d, \omega, z_\zeta) > 0$ be the constant of Theorem 6.1. Then

$$\|F'(c)\|_{\mathcal{L}} \leq \exp(BT)\,\|P\|_{H^6(Q_T)},\ \forall c \in \overline{G}.$$

In addition, the operator $F'(c)$ is Lipschitz continuous,

$$\|F'(c_1) - F'(c_2)\|_{\mathcal{L}} \leq \exp(BT) \|P\|_{H^6(Q_T)} \|c_1 - c_2\|, \ \forall c_1, c_2 \in \overline{G}.$$

Following (2.2), we also introduce the error of the level δ in the data $g(x,t)$ in (6.4). So, we assume that

$$g(x,t) = g^*(x,t) + g_\delta(x,t); \ g^*, g_\delta \in L_2(S_T), \|g_\delta\|_{L_2(S_T)} \leq \delta. \tag{6.18}$$

where $g^*(x,t)$ is the exact data and the function $g_\delta(x,t)$ represents the error in these data. To make sure that the operator F is one-to-one, we need to refer to a uniqueness theorem for our CIP. However, uniqueness results for multidimensional CIPs with single measurement data are currently known only under the assumption that at least one of initial conditions does not equal zero in the entire domain $\overline{\Omega}$, which is not our case. All these theorems were proven by the method, which was originated in 1981 simultaneously and independently by the authors of the papers [17, 18, 27]; also see, e.g. [19, 28, 29, 30] as well as sections 1.10, 1.11 of the book [11] and references cited there for some follow up publications of those authors about this method. This method is based on Carleman estimates. Although many other researchers have published about this method, we do not cite those works here, because the topic of uniqueness is not a focus of the current paper. We refer to a survey [46] for more references. Lifting the above assumption is a long standing and well known open question. Nevertheless, because of applications, it makes sense to develop numerical methods for the above CIP, regardless on the absence of proper uniqueness theorems. Therefore, we introduce Assumption 6.1.

Assumption 6.1. *The operator $F(c)$ defined in (6.17) is one-to-one.*

Theorem 6.3 follows from Theorems 3.3, 6.1 and 6.2. Note that if a function $c \in H$ is such that $c \in [1, d]$, then $c \in G$.

Theorem 6.3. *Let $\Omega \subset \mathbb{R}^3$ be a convex bounded domain with the boundary $\partial\Omega \in C^3$. Suppose that there exist functions $a(x), P(x,t), \Phi(x,t)$ satisfying conditions of Theorem 6.1. Let Assumption 6.1 and condition (6.18) hold. Let the function $v = v(x,t,c) \in H^2(Q_T)$ in (6.11) be the solution of the state problem (6.9) for the function $c \in G$. Assume that there exists the exact solution $c^* \in G, c^*(x) \in [1,d]$ of the equation $F(c^*) = 0$ for the case when in (6.18) the function g is replaced with the function g^*. Let in (6.18)*

$$\alpha = \alpha(\delta) = \delta^{2\mu}, \mu = const. \in (0, 1/4).$$

Also, let in (6.11) the function $c_{glob} \in G$ be such that

$$\|c_{glob} - c^*\| < \frac{\delta^{3\mu}}{3}.$$

Then there exists a sufficiently small number $\delta_0 = \delta_0\left(\Omega, d, \omega, z_\zeta, a, \|P\|_{H^6(Q_T)}, \mu\right) \in (0,1)$ such that for all $\delta \in (0, \delta_0)$ the neighborhood $V_{\delta^{3\mu}}(c^*)$ of the function c^* is such that $V_{\delta^{3\mu}}(c^*) \subset G$ and the functional $E_\alpha(c)$ is strongly convex in $V_{\delta^{3\mu}}(c^*)$ with the strong convexity constant $\alpha/4$. In other words,

$$\|c_1 - c_2\|^2 \leq \frac{2}{\delta^{2\mu}} \left(E'_\alpha(c_1) - E'_\alpha(c_2), c_1 - c_2\right), \ \forall c_1, c_2 \in G, \tag{6.19}$$

where $(,)$ is the scalar product in $L_2(\Omega)$ and the Fréchet derivative E'_α is calculated as in (6.15). Furthermore, there exists the unique regularized solution $c_{\alpha(\delta)}$, and $c_{\alpha(\delta)} \in V_{\delta^{3\mu}/3}(x^*)$. In addition, the gradient method of the minimization of the functional $E_\alpha(c)$, which starts at c_{glob}, converges to $c_{\alpha(\delta)}$. Furthermore, let $\xi \in (0,1)$ be an arbitrary number. Then there exists a number $\delta_1 = \delta_1\left(\Omega, d, \omega, z_\zeta, a, \|P\|_{H^6(Q_T)}, \mu, \xi\right) \in (0, \delta_0)$ such that

$$\left\|c_{\alpha(\delta)} - c^*\right\| \leq \xi \|c_{glob} - c^*\|, \ \forall \delta \in (0, \delta_1).$$

In other words, the regularized solution $c_{\alpha(\delta)}$ provides a better accuracy than the solution obtained on the first stage of our two-stage numerical procedure. Furthermore, the second relation (5.10) and (6.19) imply that

$$\left\|c - c_{\alpha(\delta)}\right\| \leq \frac{2}{\delta^{2\mu}} \left\|E'_{\alpha}(c)\right\|_{L_2(\Omega)}, \forall c \in \overline{G}. \tag{6.20}$$

Theorem 6.4 is follows from Theorems 5.1 and 6.3 as well as from Theorem 4.11.3 of [11].

Theorem 6.4. *Let conditions of Theorem 6.3 hold. Let $\|c^*\| \leq A$, where the constant A is given. Let $M_n \subset H$ be the subspace obtained after n mesh refinements as described in section 4. Also, let $V_{\delta^{3\mu}}(c^*) \cap M_1 \neq \varnothing$ and let h_n be the maximal grid step size of the subspace M_n (also, see Remark 4.1 about h_n). Let $B = B(\Omega, a, d, \omega, z_\zeta) > 0$ be the constant of Theorem 6.1 and K be the constant in (4.6). Then there exists a constant $\overline{N}_2 = \overline{N}_2 \left(\exp(BT) \|P\|_{H^6(Q_T)}\right)$ such that if*

$$h_n \leq \frac{\delta^{4\mu}}{A\overline{N}_2 K},$$

then there exists the unique minimizer c_n of the functional (6.11) on the set $G \cap M_n$, $c_n \in V_{\delta^{3\mu}}(x^*) \cap M_n$ and the following a posteriori error estimate holds

$$\left\|c_n - c_{\alpha(\delta)}\right\| \leq \frac{2}{\delta^{2\mu}} \left\|E'_{\alpha(\delta)}(c_n)\right\|_{L_2(\Omega)}. \tag{6.21}$$

The estimate (6.21) is *a posteriori* because it is obtained after the function c_n is calculated. Theorem 6.5 follows from Theorems 5.2, 5.3, 6.4 as well as from Theorem 4.11.4 of [11].

Theorem 6.5 (relaxation property of the adaptivity)**.** *Assume that conditions of Theorem 6.4 hold. Let $c_n \in V_{\delta^{3\mu}}(x^*) \cap M_n$ be the unique minimizer of the Tikhonov functional (6.11) on the set $G \cap M_n$ (Theorem 6.4). Assume that the regularized solution $c_{\alpha(\delta)} \neq c_n$, i.e. $c_{\alpha(\delta)} \notin M_n$. Let $\eta \in (0,1)$ be an arbitrary number. Then one can choose the maximal grid size $h_{n+1} = h_{n+1}\left(A, \overline{N}_2, K, \delta, z_\zeta, \mu, \eta\right) \in (0, h_n]$ of the mesh refinement number $(n+1)$ so small that (also, see Remark 4.1 about h_{n+1})*

$$\left\|c_{n+1} - c_{\alpha(\delta)}\right\| \leq \eta \left\|c_n - c_{\alpha(\delta)}\right\| \leq \frac{2\eta}{\delta^{2\mu}} \left\|E'_{\alpha(\delta)}(c_n)\right\|_{L_2(\Omega)}, \tag{6.22}$$

where the number $\overline{N}_2$ was defined in Theorem 6.4. Let $\xi \in (0,1)$ be an arbitrary number. Then there exists a sufficiently small number $\delta_0 = \delta_0\left(A, \overline{N}_2, K, \delta, z_\zeta, \xi, \mu, \eta\right) \in (0,1)$ and a decreasing sequence of maximal grid step sizes $\{h_k\}_{k=1}^{n+1}, h_k = h_k\left(A, \overline{N}_2, K, \delta, z_\zeta, \xi, \mu.\eta\right)$ such that if $\delta \in (0, \delta_0)$, then

$$\|c_{k+1} - c^*\| \leq \eta^k \left\|c_1 - c_{\alpha(\delta)}\right\| + \xi \|c_{glob} - c^*\|, k = 1, ..., n. \tag{6.23}$$

Theorem 6.6 follows from Theorems 5.4 and 6.5.

Theorem 6.6. (relaxation property of the adaptivity for local mesh refinements). *Assume that conditions of Theorem 6.5 hold. Let $\Omega = \Omega_1 \cup \Omega_2$. Suppose that mesh refinements are performed only in the subdomain Ω_2. Let $h^{(1)}$ be the maximal grid step size in Ω_1. Then there exists a sufficiently small number $\delta_0 = \delta_0\left(A, \overline{N}_2, K, \delta, z_\zeta, \xi, \mu, \eta\right) \in (0,1)$ and a decreasing sequence of maximal grid step sizes $\left\{\widetilde{h}_k\right\}_{k=1}^{n+1}, \widetilde{h}_k = \widetilde{h}_k\left(A, \overline{N}_2, K, \delta, z_\zeta, \xi, \mu, \eta\right)$ of meshes in Ω_2 such that if $\left\|\nabla c_{\alpha(\delta)}\right\|_{L_\infty(\Omega_1)}$ is so small that if*

$$\frac{2K\overline{N}_3}{\delta^{2\mu}} \left\|\nabla c_{\alpha(\delta)}\right\|_{L_\infty(\Omega_1)} h^{(1)} \leq \frac{\eta}{2} \left\|c_k - c_{\alpha(\delta)}\right\|, k = 1, ..., n \text{ and } \delta \in (0, \delta_0),$$

then (6.23) holds with the replacement of $\{h_k\}_{k=1}^{n+1}$ with $\left\{\widetilde{h}_k\right\}_{k=1}^{n+1}$. Here the number $\overline{N}_3$ depends on the same parameters as $\overline{N}_2$.

7. Mesh Refinement Recommendations and the Adaptive Algorithm.

7.1. Mesh Refinement Recommendations. Recommendations for mesh refinements are based on the theory of section 6. We now present some partly rigorous and partly heuristic considerations which lead to these recommendations. Since our considerations are partly heuristic, then both mesh refinement recommendations listed below should be verified numerically. We come back to the arguments presented in the paragraph above Theorem 5.4. To simplify the presentation, consider, for example the case (5.20), which now can be rewritten as $\nabla c_{\alpha(\delta)}(x) = \nabla c^*(x) = 0$ for $x \in \Omega_1$. A more general case of functions $c_{\alpha(\delta)}, c^*$ which change slowly in Ω_1 can be considered similarly. By (4.5) $\left(c_{\alpha(\delta)} - P_k c_{\alpha(\delta)}\right)(x) = 0$ for $x \in \Omega_1, \forall k \geq 1$. Hence, by (4.5)

$$\left\| c_{\alpha(\delta)} - P_{n+1} c_{\alpha(\delta)} \right\|_{L_2(\Omega)} = \left\| c_{\alpha(\delta)} - P_{n+1} c_{\alpha(\delta)} \right\|_{L_2(\Omega_2)} \leq K \left\| \nabla c_{\alpha(\delta)} \right\|_{L_\infty(\Omega_2)} \widetilde{h}_{n+1},$$

where $\widetilde{h}_{n+1}$ is the maximal grid step size in the subdomain Ω_2 after $n+1$ mesh refinements. Hence, using the second equality (5.11) and (5.13), we obtain

$$\left\| c_{n+1} - c_{\alpha(\delta)} \right\| \leq \frac{2K}{\delta^{2\mu}} \left\| E'_\alpha(c_{n+1}) \right\| \left\| \nabla c_{\alpha(\delta)} \right\|_{L_\infty(\Omega_2)} \widetilde{h}_{n+1}. \tag{7.1}$$

Given a function $f \in C\left(\overline{\Omega}\right)$, the main impact in the norm $\|f\|_{L_2(\Omega)}$ is provided by neighborhoods of those points $x \in \overline{\Omega}$ where the function $|f(x)|$ achieves its maximal value. Consider now formula (6.15) for the function $E'_\alpha(c)(x) \in C\left(\overline{\Omega}\right)$ (Theorem 6.1). Hence, (7.1) indicates that we should decrease the maximal grid step size $\widetilde{h}_{n+1}$ (i.e. refine mesh) in neighborhoods of those points $x \in \Omega_2$ where the function $|E'_\alpha(c_{n+1})(x)|$ achieves its maximal values. Although after n mesh refinements we know only the function $c_n \in M_n$ rather than the function $c_{n+1} \in M_{n+1}$, still, since functions c_n and c_{n+1} are sufficiently close to each other, we should likely refine mesh in neighborhoods of those points $x \in \Omega_2$ where the function $|E'_\alpha(c_n)(x)|$ achieves its maximal values. These considerations lead to mesh refinement recommendations below.

The First Mesh Refinement Recommendation. *Refine the mesh in neighborhoods of those grid points $x \in \Omega_2$ where the function $|E'_\alpha(c_n)(x)|$ attains its maximal values. Here the function $E'_\alpha(c_n)(x)$ is* given by formula *(6.15). More precisely, let $\beta_1 \in (0,1)$ be the tolerance number. Refine the mesh in such subdomains of Ω_2 where*

$$|E'_\alpha(c_n)(x)| \geq \beta_1 \max_{\overline{\Omega}_2} |E'_\alpha(c_n)(x)|. \tag{7.2}$$

To figure out the second mesh refinement recommendation, we note that by (6.15) and (6.16)

$$\left| E'_{\alpha(\delta)}(c_n)(x) \right| \leq \alpha \left(\|c_n\|_{C(\overline{\Omega})} + \|c_{glob}\|_{C(\overline{\Omega})} \right) + \|c_n\|^2_{C(\overline{\Omega})} \exp(BT) \left(\|P\|^2_{H^6(Q_T)} + \|\Phi\|^2_{H^5(Q_T)} \right).$$

Since α is small, then the second term in the right hand side of this estimate dominates. Next, since we have decided to refine the mesh in neighborhoods of those points, which deliver maximal values for the function $\left| E'_{\alpha(\delta)}(c_n)(x) \right|$, then we obtain the following mesh refinement recommendation.

Second Mesh Refinement Recommendation. *Refine the mesh in neighborhoods of those grid points $x \in \Omega_2$ where the function $c_n(x)$ attains its maximal values. More precisely, let $\beta_2 \in (0,1)$ be the tolerance number. Refine the mesh in such subdomains of Ω_2 where*

$$c_n(x) \geq \beta_2 \max_{\overline{\Omega}_2} c_n(x), \tag{7.3}$$

In fact, these two mesh refinement recommendations do not guarantee of course that the minimizer obtained on the corresponding finer mesh would be indeed more accurate than the one obtained on the coarser mesh. This is because right hand sides of formulas (7.2) and (7.3) are indicators only. Thus, numerical verifications are necessary. Nevertheless, arguments presented above in this subsection show that both recommendations are close to rigorous guarantees of accuracy improvements. As to tolerance numbers β_1 and β_2, they should be chosen numerically. Indeed, if we would choose $\beta_1, \beta_2 \approx 1$, then we would refine the mesh in too narrow regions. On the other hand, if we would choose $\beta_1, \beta_2 \approx 0$, then we would refine the mesh in almost the entire subdomain Ω_2, which is inefficient.

Remark 7.1. It was demonstrated numerically on Figure 4.28 of [11] that maximal values of the function $|E'_\alpha(c^*)(x)|$ indeed occur in places where small abnormalities are located. In other words, the gradient of the Tikhonov functional "senses" those inclusions. This can be considered as a numerical confirmation of the First Mesh Refinement Recommendation.

7.2. The adaptive algorithm. Since this algorithm was described in detail in a number of publications, see, e.g. [11, 13], we outline it only briefly here. Recall that the adaptivity is used on the second stage of our two-stage numerical procedure (section 1). On the first stage the approximately globally convergent algorithm is applied. It was proven, within the framework of the most recent Second Approximate Mathematical Model, that this algorithm delivers a good approximation for the exact solution $c^*(x)$ of the above CIP, see section 2.9 and Theorem 2.9.4 of [11]. We start the adaptivity on the same mesh on which the first stage algorithm has worked. In our experience, this mesh does not provide an improvement of the image. On each mesh we find an approximate solution of the equation $E'_\alpha(c) = 0$. Hence, by we find an approximate solution of the following equation on each mesh

$$\alpha(c - c_{glob})(x) - \int\limits_0^T (u_t\lambda_t)(x,t)\ dt = 0.$$

For each newly refined mesh we first linearly interpolate the function $c_{glob}(x)$ on it. Since this function was initially computed as a linear combination of finite elements forming the initial mesh and since all our finite elements are piecewise linear functions, then subsequent linear interpolations on finer meshes do not change the function $c_{glob}(x)$. On each mesh we iteratively update approximations c_α^n of the function $c_{\alpha(\delta)}$. To do so, we use the quasi-Newton method with the classic BFGS update formula with the limited storage [39]. Denote

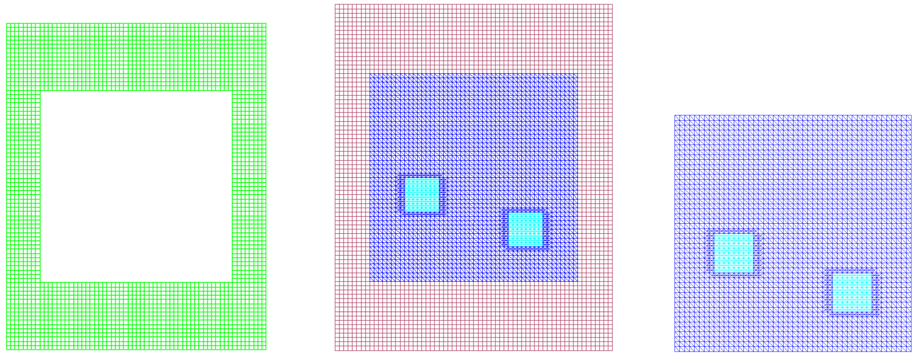
$$\varphi^n(x) = \alpha(c_\alpha^n - c_{glob})(x) - \int_0^T (v_{ht}\lambda_{ht})(x,t,c_\alpha^n)\,dt,$$

where functions $u_h(x,t,c_\alpha^n), \lambda_h(x,t,c_\alpha^n)$ are FEM solutions of state and adjoint problems (6.9), (6.10) with $c := c_\alpha^n$. We stop computing c_α^n if either $||\varphi^n||_{L_2(\Omega)} \leq 10^{-5}$ or norms $||\varphi^n||_{L_2(\Omega)}$ are stabilized. Of course, only discrete norms are considered here.

For a given mesh obtained after n mesh refinements, let c_n be the last computed function on which we have stopped. Next, we compute the function $|E'_\alpha(c_n)(x)|$ using (6.15), where $v := v_h(x,t,c_n), \lambda := \lambda_h(x,t,c_n)$. If we use both above mesh refinement recommendations, then we refine the mesh in neighborhoods of all grid points satisfying (7.2) and (7.3). In some studies, however, we use only the first recommendation. In this case we refine the mesh in neighborhoods of all grid points satisfying only (7.2).

8. Numerical Studies. We present here three numerical examples of the performance of our two-stage numerical procedure: one for computationally simulated and two for experimental data. More numerical tests of the adaptivity technique can be found in [1, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15]. In Test 1 we have used only the First Mesh Refinement Recommendation, and in Tests 2,3 we have used both recommendations. Since the numerical method of the first stage of our procedure is not a focus of this paper, and since it was described earlier in, e.g. [11, 12, 13, 14, 31, 34, 35], we do not describe it here.

8.1. Computationally simulated data. Test 1. We have made computational simulations in two dimensions. Since it is impossible to computationally solve equation (6.2) in the entire space $\mathbb{R}^2$, we have conducted data simulations in the rectangle $G = [-4,4] \times [-5,5]$. To simulate the boundary data $g(x,t)$, we have solved the forward problem by the hybrid FEM/FDM method [5] using the software package WavES [45]. To do this, we split the domain G in two subdomains $G = G_{FEM} \cup G_{FDM}$, see Figure 8.1. Here



(a) G_{FDM} (b) $G = G_{FEM} \cup G_{FDM}$ (c) $G_{FEM} = \Omega$

Fig. 8.1: The hybrid mesh (b) is a combinations of a structured mesh (a), where FDM is applied, and a mesh (c), where we use FEM, with a thin overlapping of structured elements. The solution of the inverse problem is computed in the square Ω and $c(x) = 1$ for $x \in G \diagdown \Omega$.

$G_{FEM} := \Omega = [-3,3] \times [-3,3]$ and $G_{FDM} = G \diagdown G_{FEM}$. The coefficient $c(x)$ is unknown in the domain $\Omega \subset G$ and is defined as

$$c(x) = \begin{cases} 1 & \text{in } G_{FDM}, \\ 1 + b(x) & \text{in } G_{FEM}, \\ 4 & \text{in small squares of Figure 8.1}, \end{cases} \tag{8.1}$$

where the function $b(x) \in \Omega_{FEM}$ is defined as

$$b(x) = \begin{cases} 0 & \text{for} (x_1, x_2) \in \Omega_{FEM} : -2.875 < x_1 < 0, -2.875 < x_2 < 0, \\ 0.5 \sin^2 \left(\frac{\pi x_1}{2.875} \right) \sin^2 \left(\frac{\pi x_2}{2.875} \right) & \text{otherwise.} \end{cases}$$

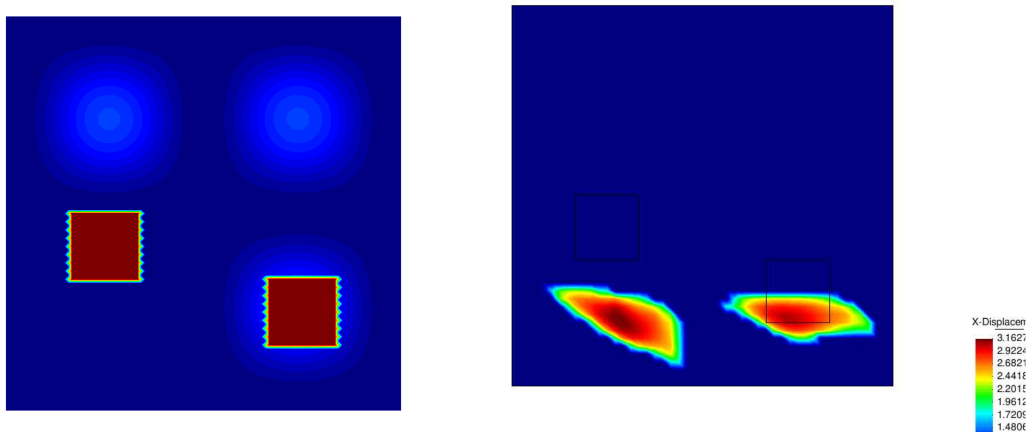
The spatial mesh consists of triangles in G_{FEM} and of squares in G_{FDM} with the grid step size $\overline{h} = 0.125$ both in overlapping regions and in G_{FDM}. There is no reason to refine mesh in G_{FDM} since $c(x) = 1$ in G_{FDM}. Let ∂G_1 and ∂G_2 be, respectively, top and bottom sides of the rectangle G and ∂G_3 be the union of vertical sides of G. We use first order absorbing boundary conditions on $\partial G \cup \partial G_2$ [21] and zero Neumann boundary condition on ∂G_3.

Let $\overline{s}$ be the upper value of the Laplace transform of the solution of our forward problem. We use this transform on the first stage of our two-stage numerical procedure. It was found that for the above domain Ω the optimal value is $\overline{s} = 7.45$. Consider the function $f(t)$,

$$f(t) = \begin{cases} 0.1 \left[\sin(\overline{s}t - \pi/2) + 1 \right], t \in [0, t_1 2\pi/\overline{s}], t_1 = 2\pi/\overline{s}, \\ 0, t \in (t_1, T], T = 17.8 t_1. \end{cases}$$

The forward problem for data simulations is

$$\begin{aligned} c(x) u_{tt} - \Delta u &= 0, \quad \text{in } G \times (0,T), \\ u(x,0) &= 0, \ u_t(x,0) = 0, \ \text{in } G, \\ \partial_n u \big|_{\partial G_1} &= f(t), \ \text{on } \partial G_1 \times (0, t_1], \\ \partial_n u \big|_{\partial G_1} &= -\partial_t u, \ \text{on } \partial G_1 \times (t_1, T), \\ \partial_n u \big|_{\partial G_2} &= -\partial_t u, \ \text{on } \partial G_2 \times (0, T), \\ \partial_n u \big|_{\partial G_3} &= 0, \ \text{on } \partial G_3 \times (0, T). \end{aligned} \tag{8.2}$$



a) exact coefficient $c(x)$ b) Coefficient c_{glob} reconstructed on the first stage

Fig. 8.2: a) Spatial distribution of the exact coefficient $c(x)$. b) Result of the performance of the approximately globally convergent algorithm (first stage). The spatial distribution of the computed coefficient c_{glob} displayed. Here $\max c_{glob}\,(x) = 3.2$, whereas $\max c\,(x) = 4$. Hence, we have 20% error in imaging of the maximal value of the function $c\,(x)$. The slowly changing part of the function $c\,(x)$, i.e. the second raw in the above definition of the function $b\,(x)$, is not imaged. Comparison with Figure a) shows that while the location of the right inclusion is imaged correctly, the left one still needs to be moved upwards. This is done on the second stage of our two-stage numerical procedure, i.e. on the adaptivity stage. On this stage we take the function $c_{glob}\,(x)$ as the starting point for the minimization of the Tikhonov functional (6.11). The second stage refines the image of the first.

The solution of this problem gives us the function $g\,(x,t) = u\mid_{S_T}$. Next, the coefficient $c\,(x)$ is "forgotten" and we apply the two-stage numerical procedure to reconstruct it from the function $g\,(x,t)$. To have noisy data, we have added the random noise to the function $g\,(x,t)$ as

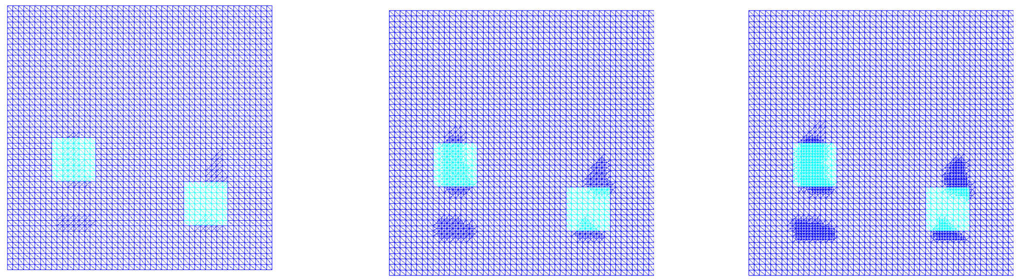
$$g_{i,j} = g\left(x^i, t^j\right)\left[1 + 0.02\alpha_j\left(g_{\max} - g_{\min}\right)\right]. \tag{8.3}$$

Here $x^i \in \partial\Omega$ and $t^j \in [0,T]$ are mesh points on $\partial\Omega$ and $[0,T]$ respectively, $g_{\min}$ and $g_{\max}$ are minimal and maximal values of the function g and $\alpha_j \in [-1,1]$ is the random variable.

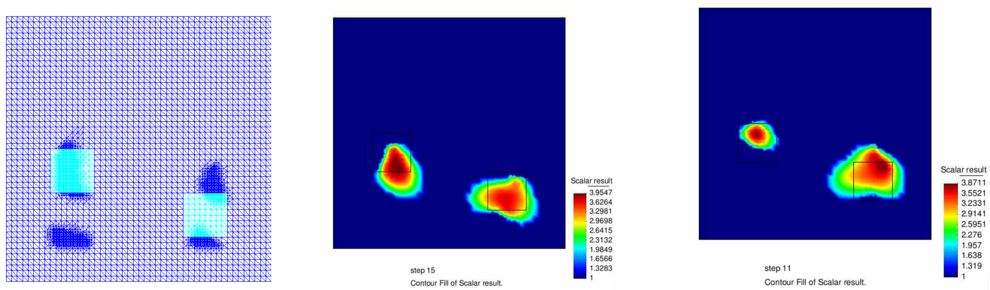
1. The approximately globally convergent stage. Since we focus on the adaptivity in this paper, we do not describe this algorithm here and refer to section 2.6.1 of [11] instead. Figure 8.2 displays the result of this stage.

2. The adaptivity stage. Since we have observed that $u\,(x,T) \approx 0$, we have not used the function $z_\zeta\,(t)$ in our computations. We now comment on the stopping criterion for mesh refinements, which we use in numerical studies of the adaptivity technique in this paper. Let c_n is the coefficient $c(x)$ calculated after n mesh refinements. In Theorems 5.2-5.4, 6.5, 6.6 the relaxation parameter η is independent on the mesh refinement number n. In practice, however, one should expect such dependence $\eta := \eta_n$. In this case the parameter η of those theorems is $\eta = \max(\eta_n)$. Then because of the relaxation property of Theorems 6.5, 6.6 as well as because of Remark 5.1, it is anticipated that numbers η_n decrease with the grow of n until the regularized solution $c_{\alpha(\delta)}$ is approximately reached. However, nothing can be guaranteed about numbers η_n as soon as the regularized solution is reached. Hence, in our computations of the adaptivity method we stopped mesh refinement process at such $n := n_0$ that $\eta_{n_0} > \eta_{n_0-1}$. If $\eta_{n_0} \approx \eta_{n_0-1}$, then we took the final solution $c_{final} := c_{n_0}$.

Figure 8.3-e),f) represents the images obtained after 4 and 5 mesh refinements, respectivelly, as well as adaptive locally refined meshes are presented on 8.3-a)-d). Comparing with Figure 8.1-c), one can observe that locations of both inclusions are imaged accurately. Recall that in each inclusion of Figure 8.1-c) $c\,(x) = 4$,

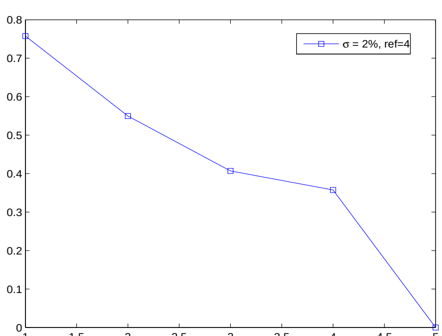


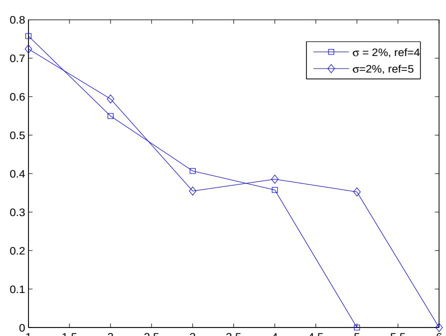
a) 4776 elements b) 5272 elements c) 6174 elements



d) 7682 elements e) $c_4(x), \max c_4(x) = 3.9$ f) $c_5(x), \max c_5(x) = 3.87$

Fig. 8.3: Adaptively refined meshes (a)–(d) and finally reconstructed images (e) and (f) on 4-th and 5-th adaptively refined meshes, respectively. On e) $\max c_4 = 3.9$ and on f) $\max c_5 = 3.87$. Reconstructed function on e) is obtained on the mesh presented on d). The mesh for the function on f) is not shown. Computational tests were performed with the noise level 2% in (8.3) and the regularization parameter $\alpha = 0.02$ in (6.11). Locations of both squares of Figure 8.3-a) as well as maximal values of the computed funtion $c_{glob}(x)$ in them are imaged accurately.





a) b)

Fig. 8.4: a) Computed relaxation property $||c_{n+1} - c_\alpha||_{L_2} \leq \eta_n ||c_n - c_\alpha||_{L_2}$ for the noise level 2% in (8.3) and the regularization parameter $\alpha = 0.02$ in (6.11). Here, $0 < \eta_n < 1$ is the small relaxation parameter obtained after n mesh refinements. Here, we take c_α on the 4-th refined mesh shown on the Figure 8.3-d). b) Comparison of the relaxation property $||c_{n+1} - c_\alpha||_{L_2} \leq \eta_n ||c_n - c_\alpha||_{L_2}$ when we take different functions c_α: on the 4-th or on the 5-th refined mesh.

see definition for $c(x)$ in (8.1) shown also on Figure 8.2-a). Therefore, maximal values of the function $c(x)$ on Figures 8.3-e),f) are also accurately imaged: the error does not exceed 3.5%.

Figure 8.4 displays the graph of the dependence of the norm $\|c_n - c_\alpha\|_{L_2(\Omega)}$ from the mesh refinement number n. By (6.22) and (6.23) these norms should decay. Since we do not exactly know what the regularized solution c_α is, we have taken $c_\alpha := c_4$ on Figure 8.4-a). On Figure 8.4-b) we have superimposed those graphs for $c_\alpha := c_4$ and $c_\alpha := c_5$. One can observe that norms $\|c_n - c_\alpha\|$ decay in the case when c_α is taken on the 4-th refined mesh. At the same time we also observe, that the relaxation property (6.23) is not fullfilled when we take c_α on the 5-th refined mesh since $\eta_3 > \eta_2$, see 8.4-b). Thus, we take the final reconstruction $c_\alpha := c_4$, the function obtained after four (4) mesh refinements.

Remark 8.1. It is well known that imaging of locations of small inclusions and maximal values of the function $c(x)$ in them is of the primary interest in applications and it is more interesting than imaging of slowly changing parts. Indeed, small inclusions can be explosives [34, 35], tumors, etc..

Remark 8.2. The above stopping criterion for mesh refinements shows that relaxation Theorems 6.5,6.6 are quite useful for computations.

8.2. Experimental data. Experimental studies were described in detail in [14, 31] as well as in Chapter 5 of [11]. Hence, we omit many details here. We point out that the main difficulty was a *huge misfit* between computationally simulated and experimental data. The latter was the case even for the free space data: the analytic solution predicted by Maxwell equations was radically different from the experimentally measured curves. This can be explained by unknown nonlinear processes in both transmitters and detectors. The same was observed for the backscattering data collected in the field, see [35] and section 6.9 of [11]. To handle this misfit, a new data pre-processing procedure was applied. This procedure has immersed experimental data in computationally simulated ones, see Figures 4 in [35] and Figures 5.3 in [11]. Naturally, this procedure has introduced a significant modeling noise in already noisy data. Nevertheless, computational results were very accurate ones, which speaks well for the robustness of our reconstruction method. The first stage of our two-stage numerical procedure was working with *blind* data (unlike the second stage). Therefore, results of at least the first stage were unbiased.

The data collection scheme is displayed on Figure 8.5. A single source of electric wave field emits pulse for only one component of the electric field, two other components were not emitted. The prism is our computational domain Ω. The outcome time resolved signal was measured at many detectors located on the bottom side of the prism. The same component of the electric field was measured as the one emitted. Since we have not measured that signal at the rest $\partial_1\Omega$ of $\partial\Omega$, we have prescribed to $\partial_1\Omega$ the same boundary conditions as ones for the uniform medium with the dielectric constant $\varepsilon_r \equiv 1$. The prism Ω is filled with a dielectric material with the dielectric constant $\varepsilon_r \approx 1$, i.e. almost the same as in the air. We point out, however, that when using the first stage of our two-stage numerical procedure, we did not use any knowledge of the dielectric constant of this prism. We have only used the fact that $\varepsilon_r = 1$ outside of this prism, see (6.1).

We have placed one dielectric inclusion inside of this prism. Inclusions were two wooden cubes, which we call below "Cube 1" and "Cube 2". Sizes of their sides were 4 cm for Cube 1 and 6 cm for Cube 2. Note that only refractive indices $n = \sqrt{\varepsilon_r}$ rather than dielectric constants can be measured directly in experiments. The goal of the first stage was to reconstruct the refractive index of the inclusion and its location. The goal of the second stage was to reconstruct all three components of inclusions: refractive indices, shapes and locations. Since only one component of the electric field was measured, we have modeled the wave propagation process via the problem (8.2) with $\varepsilon_r(x) := c(x)$, where the domain $G \subset \mathbb{R}^3$ was a prism, which was bigger than the prism Ω, see (5.8) and section 5.4 in [11] for this domain. The function $f(t)$ in (8.2) was

$$f(t) = \begin{cases} \sin(\omega t), t \in (0, 2\pi/\omega), \\ 0, t > 2\pi/\omega, \end{cases}$$

where $\omega = 14$ for Cube 1 and $\omega = 7$ for Cube 2 (see page 329 of [11] and page 26 of [14] for ω). It was only later, after the first author has conducted numerical simulations for solving the Maxwell equations [16], when we have realized that the choice of modeling by one PDE only was well justified. Indeed, it was demonstrated

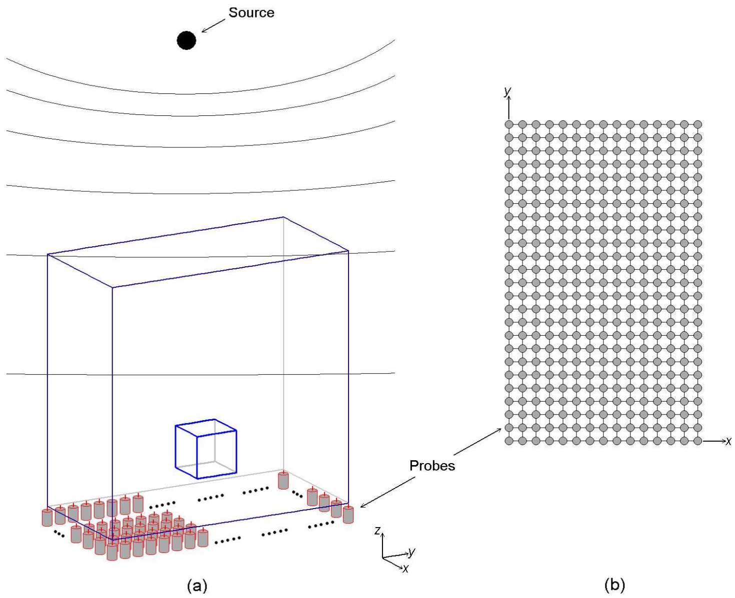


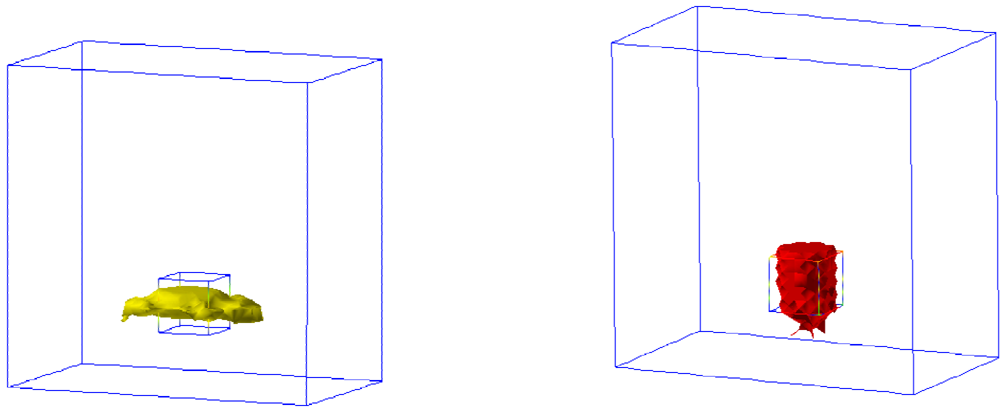
Fig. 8.5: *Schematic diagram of data collection. Original source: M. V. Klibanov, M. A. Fiddy, L. Beilina, N. Pantong and J. Schenk, Picosecond scale experimental verification of a globally convergent numerical method for a coefficient inverse problem,* Inverse Problems*, 26, 045003, doi:10.1088/0266-5611/26/4/045003, 2010. ©IOP Publishing. Reprinted with permission.*

Case number	Computed n	Directly measured n	Computational error
1 (Cube 1)	1.97	2.07	5%
2 (Cube 1)	2	2.07	3.4%
3 (Cube 1)	2.16	2.07	4.3%
4 (Cube 1)	2.19	2.07	5.8%
5 (Cube 2)	1.73	1.71	1.2%
6 (Cube 2)	1.79	1.71	4.7%

Table 8.1: Blindly computed and directly measured refractive indices n by the first stage of our two-stage numerical procedure. The error in direct measurements was 11% for cases 1-4 (Cube 1) and 3.5% for cases 5,6 (Cube 2).

computationally in [16] that the component of the electric field, which was initialized, dominates two other components. In our experiments, Cubes 1 and 2 were placed total in six different positions.

Table 8.1 summarizes results of blind study of the first stage of our two-stage numerical procedure. Because of the blind test requirement, direct measurements of refractive indices were performed by the conventional so-called "waveguide method" [41] *only after* computations of the first stage were done. Next, computational results were compared with measured ones. One can see that we had only a few percent difference with *a posteriori* directly measured refractive indices of both cubes. Furthermore, in five out of six cases this error was even less than the error in direct measurements.



a) $\max n_{glob} = \max \sqrt{\varepsilon_{r,glob}} = \max \sqrt{c_{glob}} = 1.97$ b) $\max n = \max \sqrt{\varepsilon_r} = \max \sqrt{c} = 2.05$

Fig. 8.6: Case 1 of Table 8.1 was tested by the two-stage numerical procedure. a) The computational result of the first stage. Both location of the inclusion and refractive index $n_{glob} = 1.97$ are accurately reconstructed. However, the shape of the inclusion is not reconstructed accurately. b) The computational result of the second (refinement) stage. All three components of the inclusion are very accurately reconstructed: refractive index, location and shape. Also, values of the function $\varepsilon_r(x) := c(x) = 1$ outside of the imaged inclusion are computed very accurately. Original source: L. Beilina and M.V. Klibanov, Reconstruction of dielectrics from experimental data via a hybrid globally convergent/adaptive inverse algorithm, *Inverse Problems*, 26, 125009, doi:10.1088/0266-5611/26/12/125009, 2010. ©IOP Publishing. Reprinted with permission.

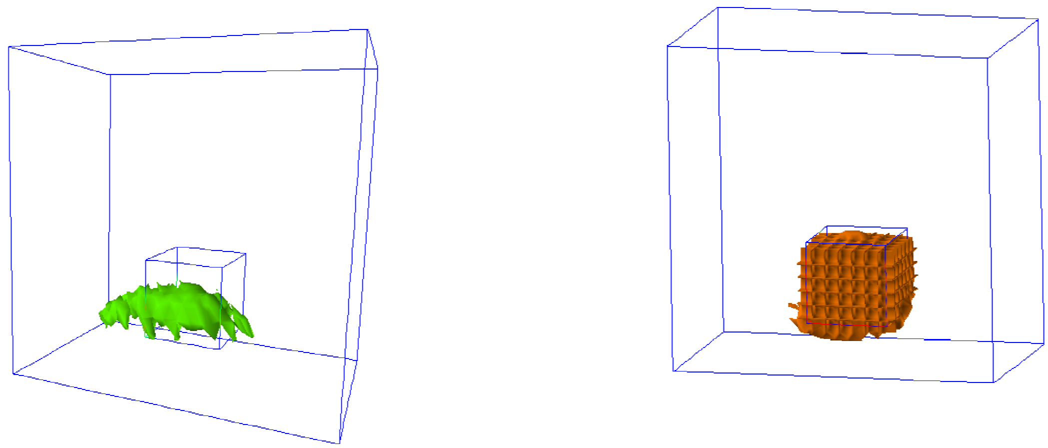
We now focus on the results which we have obtained on the second stage of our two-stage numerical procedure when applying the adaptivity.

Test 2. The two stage numerical procedure for Case 1 of Table 8.1. Figure 8.6-a) displays the result of the first stage of the two-stage numerical procedure. One can see that although the refractive index $n = 1.97$ and location of the inclusion are accurately calculated, the shape is inaccurate. The image of Figure 8.6-a) was taken as the starting point for the adaptivity technique for refinement. The result of the second stage is presented on Figure 8.6-b). One can see that all three components of the inclusion are accurately reconstructed. In addition, the values of the function $\varepsilon_r(x) := c(x) = 1$ outside of the imaged inclusion are also accurately computed.

Test 3. The two stage numerical procedure for Case 6 of Table 8.1. Figures 8.7-a) and 8.7-b) display computational results for first and second stages, respectively. The rest of comments are the same as ones for Test 2. Note that the shape is now reconstructed better than in Test 2. This can be heuristically explained as follows. The wavelength of our electromagnetic wave was $\mu = 3$ cm. Thus, the size of the side of Cube 1 is 4 cm=1.33μ. One the other hand, the size of the side of Cube 2 is 6 cm=2μ, which is larger.

Acknowledgments

This research was supported by US Army Research Laboratory and US Army Research Office grant W911NF-11-1-0399, the Swedish Research Council, the Swedish Foundation for Strategic Research (SSF) in Gothenburg Mathematical Modelling Centre (GMMC) and by the Swedish Institute, Visby Program.



a) $\max n_{glob} = \max \sqrt{\varepsilon_{r,glob}} = \max \sqrt{c_{glob}} = 1.79$ b) $\max n = \max \sqrt{\varepsilon_r} = \max \sqrt{c} = 1.73$

Fig. 8.7: Case 6 of Table 8.1 was tested for the two-stage numerical procedure. a) The computational result of the first stage. Both location of the inclusion and refractive index $n_{glob} = 1.79$ are accurately reconstructed. However, the shape of the inclusion is not reconstructed accurately. b) The computational result of the second (refinement) stage. All three components of the inclusion are accurately reconstructed: location, refractive index, and shape**.** In addition, values of the function $\varepsilon_r(x) := c(x) = 1$ outside of the imaged inclusion are computed accurately. Original source: L. Beilina and M.V. Klibanov, Reconstruction of dielectrics from experimental data via a hybrid globally convergent/adaptive inverse algorithm, *Inverse Problems*, 26, 125009, doi:10.1088/0266-5611/26/12/125009, 2010. ©IOP Publishing. Reprinted with permission.

REFERENCES

[1] M. Asadzadeh and L. Beilina, *A posteriori* error analysis in a globally convergent numerical method for a hyperbolic coefficient inverse problem, *Inverse Problems*, 26, 115007, 2010.

[2] A.B. Bakushinskii and M.Yu. Kokurin, *Iterative Methods for Approximate Solution of Inverse Problems*, Springer, New York, 2004.

[3] W. Bangerth and A. Joshi, Adaptive finite element methods for the solution of inverse problems in optical tomography, *Inverse Problems* 24, 034011, 2008.

[4] R. Becker and R. Rannacher, An optimal control approach to *a posteriori* error estimation in finite element method, *Acta Numerica*, 10, 1-102, 2001.

[5] L. Beilina, K. Samuelsson and K. Åhlander, Efficiency of a hybrid method for the wave equation. In *International Conference on Finite Element Methods*, Gakuto International Series Mathematical Sciences and Applications. Gakkotosho CO., LTD, 2001.

[6] L. Beilina and C. Johnson, A hybrid FEM/FDM method for an inverse scattering problem. In *Numerical Mathematics and Advanced Applications - ENUMATH 2001*, Springer-Verlag, Berlin, 2001.

[7] L. Beilina, Adaptive finite element/difference method for inverse elastic scattering waves, *Applied and Computational Mathematics*, 1, 158-174, 2002.

[8] L. Beilina and C. Johnson, *A posteriori* error estimation in computational inverse scattering, *Mathematical Models and Methods in Applied Sciences*, 15, 23-37, 2005.

[9] L. Beilina and C. Clason, An adaptive hybrid FEM/FDM method for an inverse scattering problem in scanning acoustic microscopy, *SIAM J. Sci. Comp.,* 28, 382-402, 2006.

[10] L. Beilina, Adaptive finite element method for a coefficient inverse problem for the Maxwell's system, *Applicable Analysis*, 90, 1461-1479, 2011.

[11] L. Beilina and M.V. Klibanov, *Approximate Global Convergence and Adaptivity for Coefficient Inverse Problems*, Springer, New York, 2012.

[12] L. Beilina and M.V. Klibanov, Synthesis of global convergence and adaptivity for a hyperbolic coefficient inverse problem in 3D, *J. Inverse and Ill-posed Problems,* 18**,** 85-132, 2010.

[13] L. Beilina and M.V. Klibanov, *A posteriori* error estimates for the adaptivity technique for the Tikhonov functional and global convergence for a coefficient inverse problem, *Inverse Problems,* 26**,** 045012, 2010.

[14] L. Beilina and M.V.Klibanov, Reconstruction of dielectrics from experimental data via a hybrid globally

convergent/adaptive inverse algorithm, *Inverse Problems*, 26, 125009, 2010.
[15] L. Beilina, M.V. Klibanov and M.Yu Kokurin, Adaptivity with relaxation for ill-posed problems and global convergence for a coefficient inverse problem, *Journal of Mathematical Sciences,* 167, 279-325, 2010.
[16] L. Beilina, Domain decomposition finite element/finite difference approach for the Maxwell's system in time domain, preprint, ISSN 1652-9715, number 2011:18; Department of Mathematical Sciences of Chalmers University and Gothenburg University, 2011.
[17] A.L. Bukhgeim and M.V. Klibanov, Uniqueness in the large of a class of multidimensional inverse problems, *Soviet Math. Doklady*, 17, 244-247, 1981.
[18] A.L. Bukhgeim, Carleman estimates for Volterra operators and uniqueness of inverse problems, in *Non-Classical Problems of Mathematical Physics*, pages 54-64, published by Computing Center of the Siberian Branch of USSR Academy of Science, Novosibirsk, 1981 (in Russian).
[19] A.L. Bukhgeim, *Introduction In The Theory of Inverse Problems*, VSP, Utrecht, The Netherlands, 2000.
[20] H.W. Engl, M. Hanke and A. Neubauer, *Regularization of Inverse Problems*, Kluwer Academic Publishers, Boston, 2000.
[21] B. Engquist and A. Majda, Absorbing boundary conditions for the numerical simulation of waves *Math. Comp.* 31, 629-651, 1977.
[22] K. Eriksson, D. Estep and C. Johnson, *Calculus in Several Dimensions*, Springer, Berlin, 2004.
[23] T. Feng, N. Yan and W. Liu, Adaptive finite element methods for the identification of distributed parameters in elliptic equation, *Advances in Computational Mathematics*, 29, 27-53, 2008.
[24] A. Griesbaum, B. Kaltenbacher and B. Vexler, Efficient computation of the Tikhonov regularization parameter by goal-oriented adaptive discretization, *Inverse Problems,* 24, 025025, 2008.
[25] B. Kaltenbacher, A. Krichner and B. Vexler, Adaptive discretizations for the choice of a Tikhonov regularization parameter in nonlinear inverse problems, *Inverse Problems*, 27, 125008, 2011.
[26] B. Kaltenbacher, A. Neubauer and O. Scherzer, *Iterative Regularization Methods for Nonlinear Ill-Posed Problems*, de Gruyter, New York, 2008.
[27] M. V. Klibanov, Uniqueness of solutions in the 'large' of some multidimensional inverse problems, in *Non-Classical Problems of Mathematical Physics*, pages 101-114, 1981, published by Computing Center of the Siberian Branch of the USSR Academy of Science, Novosibirsk (in Russian).
[28] M. V. Klibanov, Inverse problems in the 'large' and Carleman bounds, *Differential Equations*, 20, 755-760, 1984.
[29] M. V. Klibanov, Inverse problems and Carleman estimates, *Inverse Problems*, 8, 575–596, 1992.
[30] M. V. Klibanov and A. Timonov, *Carleman Estimates for Coefficient Inverse Problems and Numerical Applications*, VSP, Utrecht, 2004.
[31] M. V. Klibanov, M. A. Fiddy, L. Beilina, N. Pantong and J. Schenk, Picosecond scale experimental verification of a globally convergent numerical method for a coefficient inverse problem, *Inverse Problems*, 26, 045003, 2010.
[32] M.V. Klibanov, A.B. Bakushinskii and L. Beilina, Why a minimizer of the Tikhonov functional is closer to the exact solution than the first guess, *J. Inverse and Ill-Posed Problems*, 19, 83-105, 2011.
[33] N.A. Koshev and L. Beilina, A posteriori error estimates for the Fredholm integral equation of the first kind, accepted for publication in the book series *Springer Proceedings in Mathematics*, 2012.
[34] A. V. Kuzhuget, L. Beilina and M. V. Klibanov, Approximate global convergence and quasi-reversibility for a coefficient inverse problem with backscattered data, *Journal of Mathematical Sciences,*181, 19-49, 2012.
[35] A.V. Kuzhuget, L. Beilina, M.V. Klibanov, A. Sullivan, L. Nguyen and M.A. Fiddy, Blind experimental data collected in the field and an approximately globally convergent inverse algorithm, preprint, available online at http://www.ma.utexas.edu/mp_arc/.
[36] O. A. Ladyzhenskaya, *Boundary Value Problems of Mathematical Physics*, Springer Verlag, Berlin, 1985.
[37] J. Li, J. Xie and J. Zou, An adaptive finite element reconstruction of distributed fluxes, *Inverse Problems*, 27, 075009, 2011.
[38] M. Minoux, *Mathematical Programming: Theory and Algorithms*, Wiley and Sons, Chichester, 1986.
[39] J. Nocedal, Updating quasi-Newton matrices with limited storage, *Mathematics of Comp.*, V.35, N.151, 773–782, 1991.
[40] R. Ramlau, TIGRA- an iterative algorithm for regularizing nonlinear ill-posed problems, *Inverse Problems*, 19, 433-465, 2003.
[41] J.R. Reitz, F.J. Milford, and R.W. Christy, *Foundations of Electromagnetic Theory*, Reading, Mass.: Addison-Wesley, 1980.
[42] A. N. Tikhonov and V. Ya. Arsenin, *Solutions of Ill-Posed Problems,* Winston and Sons, Washington, DC, 1977.
[43] A.N. Tikhonov, A.V. Goncharsky, V.V. Stepanov and A.G. Yagola, *Numerical Methods for the Solution of Ill-Posed Problems*, London: Kluwer, London, 1995.
[44] A. N. Tikhonov and A.A. Samarskii, *Equations of Mathematical Physics (Dover Books on Physics)*, New York: Dover Publications, Inc., 1990.
[45] The software package WavES, available at http://waves24.com
[46] M. Yamamoto, Carleman estimates for parabolic equations and applications, *Inverse Problems,* 25**,** 123013, 2009.