Why a minimizer of the Tikhonov functional is closer to the exact solution than the first guess

Michael V. Klibanov*, Anatoly B. Bakushinsky◇ and Larisa Beilina△
*Department of Mathematics and Statistics
University of North Carolina at Charlotte, Charlotte, NC 28223, USA
◇Institute for System Analysis of The Russian Academy of Science,
Prospect 60 letya Oktyabrya 9 , 117312, Moscow, Russia
△Department of Mathematical Sciences
Chalmers University and Gothenburg University
SE-42196, Gothenburg, Sweden
E-mails: mklibanv@uncc.edu, bakush@isa.ru, larisa@chalmers.se

September 27, 2010

Abstract

Suppose that a uniqueness theorem is valid for an ill-posed problem. It is shown then that the distance between the exact solution and terms of a minimizing sequence of the Tikhonov functional is less than the distance between the exact solution and the first guess. Unlike the classical case when the regularization parameter tends to zero, only a single value of this parameter is used. Indeed, the latter is always the case in computations. Next, this result is applied to a specific Coefficient Inverse Problem. A uniqueness theorem for this problem is based on the method of Carleman estimates. In particular, the importance of obtaining an accurate first approximation for the correct solution follows from one of theorems. The latter points towards the importance of the development of globally convergent numerical methods as opposed to conventional locally convergent ones. A numerical example is presented

1 Introduction

The classical regularization theory for nonlinear ill-posed problems guarantees convergence of regularized solutions to the exact one only when both the level of the error δ and the regularization parameter $\alpha(\delta)$ tend to zero, see, e.g. [10,18]. However, in the computational practice one always works with a **single** value of δ and a **single** value of $\alpha(\delta)$. Let x^* be the ideal exact solution of an ill-posed problem for the case of ideal noiseless data. The existence of such a solution should be assumed by one of Tikhonov principles [18]. Suppose that x^* is unique. The Tikhonov regularization term usually has the form $\alpha(\delta)\|x-x_0\|^2$, where x_0 is counted as the first guess for x^* and $\|\cdot\|$ is the norm in a certain space. Let $\{x_n\}_{n=1}^{\infty}$ be a minimizing sequence for the Tikhonov functional. In the case when the existence of a minimizer of this functional is guaranteed (e.g. the finite dimensional case), we replace this sequence with a minimizer x_α, which is traditionally called a *regularized solution* (one might have several such solutions).

The goal of this paper is to address the following questions: *Why it is usually observed computationally that for a proper* **single** *value of* $\alpha(\delta)$ *one has* $\|x_n - x^*\| < \|x_0 - x^*\|$ *for sufficiently large* n*? In particular, if* x_α *exists, then why the computational observation is that* $\|x_\alpha - x^*\| < \|x_0 - x^*\|$*? In other words, why the regularized solution is usually more accurate than the first guess even for a* **single** *value of* $\alpha(\delta)$?

Indeed, the regularization theory guarantees that this is true only in the limiting sense when both δ and $\alpha(\delta)$ tend to zero. These questions were not addressed in the regularization theory. We first prove a general Theorem 2, which addresses the above questions in an "abstract" form of Functional Analysis. In accordance to this theorem the short answer on these questions is this: *If a uniqueness theorem holds for an ill-posed problem, then the above accuracy improvement takes place.* Usually it is not easy to "project" an abstract theory on a specific problem. Thus, the major effort of this paper is focused on the demonstration on how our Theorem 2 works for a specific Coefficient Inverse Problem for a hyperbolic PDE. Also, we present a numerical result for this problem.

We show that two factors are important for the above observation: (1) the uniqueness theorem for the original problem, and (2) the availability of a good first guess. The latter means that x_0 should be chosen in such a way that the distance $\|x_0 - x^*\|$ would be sufficiently small. In other words, one needs to obtain a good first approximation for the exact solution, from which subsequent iterations would start from. These iterations would refine x_0. However, in the majority of applications it is unrealistic to assume that such an approximation is available. Hence, one needs to develop a *globally convergent* numerical method for the corresponding ill-posed problem. This means that one needs to work out such a numerical method which would provide a guaranteed good approximation for x^* without *a priori* knowledge of a small neighborhood of x^*.

To apply our Theorem 2 to a Coefficient Inverse Problem (CIP) for a hyperbolic PDE, we use a uniqueness theorem for this CIP, which was proved via the method of Carleman estimates [7,11-15]. Note that the technique of [7,11-15] enables one to prove uniqueness theorems for a wide variety of CIPs. A similar conclusion about the importance of the uniqueness theorem and, therefore, of the modulus of the continuity of the inverse operator on a certain compact set, was drawn in [9]. However, norms $\|x_\alpha - x^*\|$ and $\|x_0 - x^*\|$ were not compared in this work, since $x_0 = 0$ in [9]. Only the case when both δ and $\alpha(\delta)$ tend to zero was considered in [9], whereas these parameters are fixed in our case.

We now explain our motivation for this paper. The *crucial* question about a method of choice of a good first approximation for the exact solution is left outside of the classical regularization theory. This is because such a choice depends on a specific problem at hands.

On the other hand, since the Tikhonov regularization functional usually has many local minima and ravines, a successful first guess is the *true key* for obtaining an accurate solution. Recently the first and third authors have developed a globally convergent numerical method for a Coefficient Inverse Problem (CIP) for a hyperbolic PDE [2-6]. Furthermore, this method was confirmed in [16] via a surprisingly very accurate imaging results using *blind* experimental data.

It is well known that CIPs are very challenging ones due to their nonlinearity and ill-posedness combined. So, it is inevitable that an approximation was made in [2] via the truncation of high values of the positive parameter s of the Laplace transform. In fact, this approximation is similar with the truncation of high frequencies, which is routinely done in engineering. It was discovered later in [3-6] that although the globally convergent method of [2,3] provides a guaranteed good approximation for the exact solution, it needs to be refined due to that approximation. On the other hand, since the *main input* for any locally convergent method is a good first approximation for the exact solution, then a locally convergent technique was chosen for the refinement stage. The resulting two-stage numerical procedure has consistently demonstrated a very good performance [2-6], including the case of experimental data [5]. So, since only a single value of the regularization parameter was used in these publications, then this naturally has motivated us to address the above questions.

We also wish to point out that the assumption about *a priori* known upper estimate of the level of the error δ is not necessary true in applications. Indeed, a huge discrepancy between experimental and computationally simulated data was observed in [5] and [16], see Figure 2 in [5] and Figures 3,4 in [16]. This discrepancy was successfully addressed via new data pre-processing procedures described in [5,16]. Therefore, it is unclear what kind of δ the experimental data of [5,16] actually have. Nevertheless the notion of *a priori* knowledge of δ is quite useful for qualitative explanations of computational results for ill-posed problems.

In section 2 we present the general scheme. In section 3 we apply it to a CIP. In section 4 we present a numerical example of the above mentioned two-stage numerical procedure solely for the illustration purpose.

2 The General Scheme

For simplicity we consider only real valued Hilbert spaces, although our results can be extended to Banah spaces. Let H, H_1 and H_2 be real valued Hilbert spaces with norms $\|\cdot\|, \|\cdot\|_1$ and $\|\cdot\|_2$ respectively. We assume that $H_1 \subset H, \|x\| \leq \|x\|_1, \forall x \in H_1$,the set H_1 is dense in H and any bounded set in H_1 is a compact set in H. Let $W \subset H$ be an open set and $G = \overline{W}$ be its closure. Let $F : G \to H_2$ be a continuous operator, not necessarily linear. Consider the equation

$$F(x) = y, x \in G. \tag{1}$$

The right hand side of (1) might be given with an error. The existence of a solution of this equation is not guaranteed. However, in accordance with one of Tikhonov principles for ill-posed problems, we assume that there exists an ideal exact solution x^* of (1) with an ideal errorless data y^* and this solution is unique,

$$F\left(x^*\right) = y^*. \tag{2}$$

We assume that

$$\left\| y - y^* \right\|_2 \leq \delta, \tag{3}$$

where the sufficiently small number $\delta > 0$ characterizes the error in the data. Since it is unlikely that one can get a better accuracy of the solution than $O\left(\delta\right)$ in the asymptotic sense as $\delta \to 0$, then it is usually acceptable that all other parameters involved in the regularization process are much larger than δ. For example, let the number $\gamma \in (0,1)$. Since $\lim_{\delta \to 0}\left(\delta^\gamma / \delta\right) = \infty$, then there exists a sufficiently small number $\delta_0\left(\gamma\right) \in (0,1)$ such that

$$\delta^\gamma > \delta, \forall \delta \in \left(0, \delta_0\left(\gamma\right)\right). \tag{4}$$

Hence, it would be acceptable if one would choose $\alpha\left(\delta\right) = \delta^\gamma$, see (9) below.

To solve the problem (1), consider the Tikhonov regularization functional with the regularization parameter α,

$$J_\alpha\left(x\right) = \left\| F\left(x\right) - y \right\|_2^2 + \alpha \left\| x - x_0 \right\|_1^2, x \in G \cap H_1. \tag{5}$$

It is natural to treat x_0 here as the starting point for iterations, i.e. the first guess for the solution. We are interested in the question of the minimization of this functional for a fixed value of α. It is not straightforward to prove the existence of the minimizer, unless some additional conditions would be imposed on the operator F, see subsection 3.3 for a possible one. Hence, all what we can consider is a minimizing sequence. Let $m = \inf_{G \cap H_1} J_\alpha\left(x\right).$ Then

$$m \leq J_\alpha\left(x^*\right). \tag{6}$$

Usually $m < J_\alpha\left(x^*\right)$. However we do not use the latter assumption here.

2.1 The case $\dim H = \infty$

Let $\dim H = \infty$. In this case we cannot prove existence of a minimizer of the functional J_α, i.e. we cannot prove the existence of a regularized solution. Hence, we work now with the minimizing sequence. It follows from (5), (6) that there exists a sequence $\left\{x_n\right\}_{n=1}^\infty \subset G \cap H_1$ such that

$$m \leq J_\alpha\left(x_n\right) \leq \delta^2 + \alpha \left\| x^* - x_0 \right\|_1^2 \text{ and } \lim_{n \to \infty} J_\alpha\left(x_n\right) = m. \tag{7}$$

By (3)-(7)

$$\left\| x_n \right\|_1 \leq \left(\frac{\delta^2}{\alpha} + \left\| x^* - x_0 \right\|_1^2 \right)^{1/2} + \left\| x_0 \right\|_1. \tag{8}$$

It is more convenient to work with a lesser number of parameters. So, we assume that $\alpha = \alpha\left(\delta\right)$. To specify this dependence, we note that we want the right hand side of (8) to be bounded as $\delta \to 0$. So, using (4), we assume that

$$\alpha\left(\delta\right) = \delta^{2\mu}, \text{ for a } \mu \in (0,1). \tag{9}$$

Hence, it follows from (8) and (9) that $\{x_n\}_{n=1}^{\infty} \subset K(\delta, x_0)$, where $K(\delta, x_0)$ is a precompact set in H defined as

$$K(\delta, x_0) = \left\{ x \in H_1 : \|x\|_1 \leq \sqrt{\delta^{2(1-\mu)} + \|x^* - x_0\|_1^2} + \|x_0\|_1 \right\}. \tag{10}$$

Note that the sequence $\{x_n\}_{n=1}^{\infty}$ depends on δ. The set $K(\delta, x_0)$ is not closed in H. Let $\overline{K} = \overline{K}(\delta, x_0)$ be its closure. Hence, $\overline{K}$ is a closed compact set in H.

Remark 1. We have introduced the specific dependence $\alpha := \alpha(\delta) = \delta^{2\mu}$ in (9) for the sake of definiteness only. In fact, in the theory below one can consider many other dependencies $\alpha := \alpha(\delta)$, as long as $\alpha(\delta) >> \delta$ for sufficiently small values of δ and $\lim_{\delta \to 0} \alpha(\delta) = 0$.

Let B_1 and B_2 be two Banah spaces, $P \subset B_1$ be a closed compact and $A : P \to B_2$ be a continuous operator. We remind that there exists the modulus of the continuity of this operator on P. In other words, there exists a function $\omega(z)$ of the real variable $z \geq 0$ such that

$$\begin{aligned} \omega(z) &\geq 0, \omega(z_1) \leq \omega(z_2), \text{if } z_1 \leq z_2, \lim_{z \to 0^+} \omega(z) = 0, & (11) \\ \|A(x_1) - A(x_2)\|_{B_2} &\leq \omega\left(\|x_1 - x_2\|_{B_1}\right), \forall x_1, x_2 \in P. & (12) \end{aligned}$$

We also remind a theorem of A.N. Tikhonov, which is one of back bones of the theory of ill-posed problems:

Theorem 1 (A.N. Tikhonov, [18]). *Let B_1 and B_2 be two Banah spaces, $P \subset B_1$ be a closed compact set and $A : P \to B_2$ be a continuous one-to-one operator. Let $Q = A(P)$. Then the operator $A^{-1} : Q \to P$ is continuous.*

We now prove

Theorem 2. *Let the above operator $F : G \to H_2$ be continuous and one-to-one. Suppose that (9) holds. Let $\{x_n\}_{n=1}^{\infty} \subset K(\delta, x_0) \subseteq \overline{K}$ be a minimizing sequence of the functional (5) satisfying (7). Let $\xi \in (0,1)$ be an arbitrary number. Then there exists a sufficiently small number $\delta_0 = \delta_0(\xi) \in (0,1)$ such that for all $\delta \in (0, \delta_0)$ the following inequality holds*

$$\|x_n - x^*\| \leq \xi \|x_0 - x^*\|_1, \forall n. \tag{13}$$

In particular, if a regularized solution $x_{\alpha(\delta)}$ exists for a certain $\delta \in (0, \delta_0)$, i.e. $J_{\alpha(\delta)}\left(x_{\alpha(\delta)}\right) = m(\delta)$, then

$$\left\|x_{\alpha(\delta)} - x^*\right\| \leq \xi \|x_0 - x^*\|_1. \tag{14}$$

Proof. By (5), (7) and (9)

$$\|F(x_n) - y\|_2 \leq \sqrt{\delta^2 + \alpha \|x_0 - x^*\|_1^2} = \sqrt{\delta^2 + \delta^{2\mu} \|x_0 - x^*\|_1^2}.$$

Hence,

$$\begin{aligned} \|F(x_n) - F(x^*)\|_2 &= \|(F(x_n) - y) + (y - F(x^*))\|_2 \\ &= \|(F(x_n) - y) + (y - y^*)\|_2 \\ &\leq \|F(x_n) - y\|_2 + \|y - y^*\|_2 \leq \sqrt{\delta^2 + \delta^{2\mu} \|x^* - x_0\|_1^2} + \delta, \end{aligned} \tag{15}$$

where we have used (3). By Theorem 1 there exists the modulus of the continuity $\omega_F(z)$ of the operator $F^{-1}: F\left(\overline{K}\right) \to \overline{K}$. It follows from (15) that

$$\|x_n - x^*\| \le \omega_F\left(\sqrt{\delta^2 + \delta^{2\mu}\|x_0 - x^*\|_1^2} + \delta\right). \tag{16}$$

Consider an arbitrary $\xi \in (0,1)$. It follows from (11) that one can choose a number $\delta_0 = \delta_0(\xi)$ so small that

$$\omega_F\left(\sqrt{\delta^2 + \delta^{2\mu}\|x^* - x_0\|_1^2} + \delta\right) \le \xi \|x_0 - x^*\|_1, \forall \delta \in (0, \delta_0). \tag{17}$$

The estimate (13) follows from (16) and (17). It is obvious that in the case of the existence of the regularized solution, one can replace in (16) x_n with $x_{\alpha(\delta)}$. $\square$

2.2 Discussion of Theorem 2

One can see from (16) and (17) that two main factors defining the accuracy of the reconstruction of points x_n, as well as of the regularized solution $x_{\alpha(\delta)}$ (in the case when it exists) are: **(a)** the level of the error δ, and **(b)** the accuracy of the first approximation x_0 for the exact solution, i.e. the norm $\|x^* - x_0\|_1$. However, since in applications one always works with a fixed level of error δ, then the second factor is more important than the first.

In addition, regardless on the existence of the infimum value $m(\delta)$ of the functional $J_{\alpha(\delta)}$, it is usually impossible in practical computations to construct the above minimizing sequence $\{x_n\}_{n=1}^{\infty}$. This is because the functional $J_{\alpha(\delta)}$ usually features the problem of multiple local minima and ravines. The only case when an effective construction of that sequence might be possible is when one can figure out such a first guess x_0 for the exact solution x^*, which would be located sufficiently close to x^*, i.e. the norm $\|x_0 - x^*\|_1$ should be sufficiently small. Indeed, it was proven in [6] that, under some additional conditions, the functional $J_{\alpha(\delta)}$ is strongly convex in a small neighborhood of $x_{\alpha(\delta)}$ (it was assumed in [6] that the regularized solution exists). Furthermore, in the framework of [6] x^* belongs to this neighborhood, provided that x_0 is sufficiently close to x^*, see subsection 3.3 for some details. Hence, local minima and ravines do not occur in that neighborhood. In particular, the latter means that the two stage procedure of [2-6] converges *globally* to the exact solution. Indeed, not only the first stage provides a guaranteed good approximation for that solution, but also, by Theorems 7 and 8 (subsection 3.3), the second stage provides a guaranteed refinement and the gradient method, which starts from that first approximation, converges to the refined solution.

In summary, the discussion of the above paragraph points towards the *fundamental* importance of a proper choice of a good first guess for the exact solution, i.e. towards the importance of developments of globally convergent numerical methods, as opposed to conventional locally convergent ones.

Another inconvenience of Theorem 2 is that it is difficult to work with a stronger norm $\|\cdot\|_1$ in practical computations. Indeed, this norm is used only to ensure that $K(\delta, x_0)$ is a compact set in H. However, the most convenient norm in practical computations is the L_2-norm. Hence, when using the norm $\|\cdot\|_1$, one should set $\|\cdot\|_1 := \|\cdot\|_{H^p}, p \geq 1$. At the same time, convergence of the sequence $\{x_n\}_{n=1}^{\infty}$ will occur in the L_2-norm. This is clearly inconvenient in practical computations. Hence, finite dimensional spaces of standard piecewise linear finite elements were used in computations of [2-6] for both forward and inverse problems. Note that all norms in such spaces are equivalent. Still, the number of these finite elements should not be exceedingly large. Indeed, if it would, then that finite dimensional space would effectively become an infinitely dimensional one, and the L_2-norm would not be *factually* equivalent to a stronger norm. This is why local mesh refinements were used in [2-6], which is an opposite to a uniformly fine mesh. In other words, the so-called Adaptive Finite Element method (adaptivity below for brevity) was used on the second stage of the two-stage numerical procedure of [2-6]. While the adaptivity is well known for classical forward problems for PDEs, the first publication for the case of a CIP was in [1].

3 Example: a Coefficient Inverse Problem for a Hyperbolic PDE

We now consider a CIP for which a globally convergent two-stage numerical procedure was developed in [2-6]. Recall that on the first stage the globally convergent numerical method of [2,3] provides a guaranteed good first approximation for the exact solution. This approximation is *the key input* for any subsequent locally convergent method. Next, a locally convergent adaptivity technique refines the approximation obtained on the first stage. So, based on Theorem 2, we provide in this section an explanation on why this refinement became possible.

3.1 A Coefficient Inverse Problem and its uniqueness

Let $\Omega \subset \mathbb{R}^k, k = 2, 3$ be a bounded domain with its boundary $\partial\Omega \in C^3$. Consider the function $c(x)$ satisfying the following conditions

$$c(x) \in [1, d] \text{ in } \Omega, d = const. > 1, c(x) = 1 \text{ for } x \in \mathbb{R}^k \diagdown \Omega, \tag{18}$$
$$c(x) \in C^2\left(\mathbb{R}^k\right). \tag{19}$$

Let the point $x_0 \notin \overline{\Omega}$. Consider the solution of the following Cauchy problem

$$c(x) u_{tt} = \Delta u \text{ in } \mathbb{R}^k \times (0, \infty), \tag{20}$$
$$u(x, 0) = 0, u_t(x, 0) = \delta(x - x_0). \tag{21}$$

Equation (20) governs propagation of both acoustic and electromagnetic waves. In the acoustical case $1/\sqrt{c(x)}$ is the sound speed. In the 2-D case of EM waves propagation, the dimensionless coefficient $c(x) = \varepsilon_r(x)$, where $\varepsilon_r(x)$ is the relative dielectric function of the medium, see [8], where this equation was derived from the Maxwell equations in the 2-D case. It should be also pointed out that although equation (18) cannot be derived from the Maxwell's system in the 3-d case for $\varepsilon_r(x) \neq const.$, still it was successfully applied to the imaging from experimental data, and a very good accuracy of reconstruction was consistently observed [5,16], including even the case of *blind* study in [16]. We consider the following

Inverse Problem 1. *Suppose that the coefficient $c(x)$ satisfies (18) and (19), where the number $d > 1$ is given. Assume that the function $c(x)$ is unknown in the domain Ω. Determine the function $c(x)$ for $x \in \Omega$, assuming that the following function $g(x,t)$ is known for a single source position $x_0 \notin \overline{\Omega}$*

$$u(x,t) = g(x,t), \forall (x,t) \in \partial\Omega \times (0,\infty). \tag{22}$$

Since the function $c(x) = 1$ outside of the domain Ω, then we can consider (20)-(22) as the initial boundary value problem in the domain $\left(\mathbb{R}^k \diagdown \Omega\right) \times (0,\infty)$. Hence, the function $u(x,t)$ can be uniquely determined in this domain. Hence, the following function $p(x,t)$ is known

$$\partial_n u(x,t) = p(x,t), \forall (x,t) \in \partial\Omega \times (0,\infty), \tag{23}$$

where n is the unit outward normal vector at $\partial\Omega$.

This is an inverse problem with the single measurement data, since the initial condition is not varied here. This is because only a single source location is involved. The case of single measurement is the most economical one, as opposed to the case of infinitely many measurements. It is yet unknown how to prove uniqueness theorem for this inverse problem. There exists a general method of proofs of such theorems for a broad class of CIPs with single measurement data for many PDEs. This method is based on Carleman estimates, it was first introduced in [7,11] (also, see [12-15]) and became quite popular since then, see, e.g. the recent survey [19] with many references. In fact, currently this is the single method enabling proofs of uniqueness theorems for CIPs with the single measurement data. However, this technique cannot handle the case of, e.g. the $\delta-$ function in the initial condition, as it is in (21). Instead, it can handle only the case when at least one initial condition is not zero in the entire domain of interest. We use one of results of [13,14] in this section.

Fix a bounded domain $\Omega_1 \subset \mathbb{R}^k$ with $\partial\Omega_1 \in C^3$ such that $\Omega \subset\subset \Omega_1$ and $x_0 \in \Omega_1 \diagdown \overline{\Omega}$. Consider a small neighborhood $N(\partial\Omega_1)$ of the boundary $\partial\Omega_1$ such that $x_0 \notin \overline{N}(\partial\Omega_1)$. Let the function $\chi(x) \in C^\infty\left(\overline{\Omega}_1\right)$ be such that

$$\chi(x) \geq 0, \chi(x) = \begin{cases} 1 \text{ for } x \in \Omega_1 \diagdown N(\partial\Omega_1), \\ 0 \text{ for } x \in (\partial\Omega_1). \end{cases} \tag{24}$$

The existence of such functions is known from the Real Analysis course. Let $\varepsilon > 0$ be a sufficiently small number. Consider the following approximation of the function $\delta(x - x_0)$ in the distributions sense

$$f_\varepsilon(x - x_0) = C(\varepsilon, \chi, \Omega_1) \exp\left(-\frac{|x - x_0|^2}{\varepsilon^2}\right) \chi(x), \tag{25}$$

where the constant $C(\varepsilon, \chi, \Omega_1) > 0$ is chosen such that

$$\int\limits_{\Omega_1} f_\varepsilon(x - x_0)\, dx = 1. \tag{26}$$

We approximate the function u via replacing initial conditions in (21) with

$$u_\varepsilon(x, 0) = 0, \partial_t u_\varepsilon(x, 0) = f_\varepsilon(x - x_0). \tag{27}$$

Let $\overline{u}_\varepsilon(x, t)$ be the solution of the Cauchy problem (20), (27) such that

$\overline{u}_\varepsilon(x, t) \in H^2(\Phi \times (0, T))$ for any bounded domain $\Phi \subset \mathbb{R}^k$. It follows from the discussion in §2 of Chapter 4 of [17] that such a solution exists, is unique and coincides with the "target" solution $u_\varepsilon(x, t)$ of the Cauchy problem (20), (27), $\overline{u}_\varepsilon(x, t) \equiv u_\varepsilon(x, t)$. Consider a number $T > 0$, which we will define later. Then due to a finite speed of wave propagation (Theorem 2.2 in §2 of Chapter 4 of [17]) there exists a domain $\Omega_2 = \Omega_2(T)$ with $\partial\Omega_2 \in C^3$ such that $\Omega \subset\subset \Omega_1 \subset\subset \Omega_2$

$$u_\varepsilon \mid_{\partial\Omega_2 \times (0,T)} = 0 \tag{28}$$

and

$$u_\varepsilon(x, t) = 0 \text{ in } \left(\mathbb{R}^k \diagdown \Omega_2\right) \times (0, T). \tag{29}$$

Therefore, we can consider the function $u_\varepsilon(x, t)$ as the solution of the initial boundary value problem (20), (27), (28) in the domain $Q_T = \Omega_2 \times (0, T)$. Hence, repeating arguments given after the proof of Theorem 4.1 in §4 of Chapter 4 of [17], we obtain the following

Lemma 1. *Let the function $c(x)$ satisfies conditions (18), (19) and, in addition, let $c \in C^5\left(\mathbb{R}^k\right)$. Let domains Ω, Ω_1 satisfy above conditions. Let the function $f_\varepsilon(x - x_0)$ satisfies (25), (26), where the function $\chi(x)$ satisfies (24). Let the number $T > 0$. Let the domain $\Omega_2 = \Omega_2(T)$ with $\partial\Omega_2 \in C^2$ be such that (29) is true for the solution $u_\varepsilon(x, t) \in H^2(\Phi \times (0, T))$ (for any bounded domain $\Phi \subset \mathbb{R}^k$) of the Cauchy problem (20), (27). Then in fact the function $u_\varepsilon \in H^6(Q_T)$. Hence, since $k = 2, 3$, then by the embedding theorem $u_\varepsilon \in C^3\left(\overline{Q}_T\right)$.*

We need the smoothness condition $u_\varepsilon \in C^3\left(\overline{Q}_T\right)$ of this lemma because the uniqueness Theorem 3 requires it.

Theorem 3 [13,14]**.** *Assume that the function $f_\varepsilon(x - x_0)$ satisfies (25), (26) and the function $c(x)$ satisfies conditions (18), (19). In addition, assume that*

$$\frac{1}{2} + (x, \nabla c(x)) \geq \overline{c} = const. > 0 \text{ in } \overline{\Omega}. \tag{30}$$

Suppose that boundary functions $h_0(x, t), h_1(x, t)$ are given for $(x, t) \in \partial\Omega \times (0, T_1)$, where the number T_1 is defined in the next sentence. Then one can choose such a sufficiently large number $T_1 = T_1(\Omega, \overline{c}, d) > 0$ that for any $T \geq T_1$ there exists at most one pair of functions (c, v) such that $v \in C^3\left(\overline{\Omega} \times [0, T]\right)$ and

$$c(x)\, v_{tt} = \Delta v \text{ in } \Omega \times (0, T),$$

$$v(x,0)=0, v_t(x,0)=f_\varepsilon(x-x_0),$$
$$v\mid_{\partial\Omega\times(0,T)}=h_0(x,t), \partial_n v\mid_{\partial\Omega\times(0,T)}=h_1(x,t).$$

To apply this theorem to our specific case, we assume that the following functions $g_\varepsilon, p_\varepsilon$ are given instead of functions g and p in (22), (23)

$$u_\varepsilon\mid_{\partial\Omega\times(0,T)}=g_\varepsilon(x,t), \partial_n u_\varepsilon\mid_{\partial\Omega\times(0,T)}=p_\varepsilon(x,t). \tag{31}$$

Hence, using Theorem 3 and Lemma 1, we obtain the following uniqueness result

Theorem 4. *Suppose that the function $c(x)$ satisfies conditions (18), (19), $c\in C^5(\mathbb{R}^k)$ and also let the condition (30) be true. Assume that all other conditions of Lemma 1 are satisfied, except that now we assume that T is an arbitrary number such that $T\geq T_1=T_1(\Omega,\overline{c},d)$ where T_1 was defined Theorem 3. Then there exists at most one pair of functions (c,u_ε) satisfying (20), (27), (31).*

Since in Theorem 4 we have actually replaced $\delta(x-x_0)$ in (21) with the function $f_\varepsilon(x-x_0)$, then we cannot apply Theorem 2 directly to the above inverse problem. Instead, we formulate

Inverse Problem 2. *Suppose that the coefficient $c(x)$ satisfies (18), (19), (30), where the numbers $d,\overline{c}>1$ are given. In addition, let the function $c\in C^5(\mathbb{R}^k)$. Let the function $f_\varepsilon(x-x_0)$ satisfies conditions of Lemma 1. Let $T_1=T_1(\Omega,\overline{c},d)>0$ be the number chosen in Theorems 3, and $T\geq T_1$ be an arbitrary number. Consider the solution $u_\varepsilon(x,t)\in H^2(\Phi\times(0,T))$ (for any bounded domain $\Phi\subset\mathbb{R}^k$) of the following Cauchy problem*

$$c(x)\partial_t^2 u_\varepsilon = \Delta u_\varepsilon \text{ in } \mathbb{R}^k\times(0,T), \tag{32}$$
$$u_\varepsilon(x,0) = 0, \partial_t u_\varepsilon(x,0)=f_\varepsilon(x-x_0). \tag{33}$$

Assume that the coefficient $c(x)$ is unknown in the domain Ω. Determine $c(x)$ for $x\in\Omega$, assuming that the function $u_\varepsilon(x,t)$ is given for $(x,t)\in\partial\Omega\times(0,T)$.

For the sake of completeness we formulate now

Theorem 5. *Assume that conditions of Inverse Problem 2 are satisfied. Then there exists at most one solution (c,u_ε) of this problem.*

Proof. By Lemma 1 the solution of the problem (32), (33) $u_\varepsilon\in C^3(\overline{\Omega}\times[0,T])$. Since the function $c(x)=1$ in $\mathbb{R}^k\diagdown\Omega$, then, as it was shown above, one can uniquely determine the function $p_\varepsilon(x,t)=\partial_n u_\varepsilon\mid_{\partial\Omega\times(0,T)}$. Hence, Theorem 4 leads to the desired result. $\square$

3.2 Using Theorems 2-5

Following (5), we now construct the regularization functional $J_\alpha(c)$. Choose a number $T\geq T_1(\Omega,\overline{c},d)>0$, where T_1 was chosen in Theorems 3,4. Let $\Omega_2=\Omega_2(T)$ be the domain chosen in Lemma 1. Because of (29), it is sufficient to work with domains Ω,Ω_2 rather than with Ω and $\mathbb{R}^k$. In doing so, we must make sure that all functions $\{c_n\}_{n=1}^\infty$, which will replace the sequence $\{x_n\}_{n=1}^\infty$ in (7), are such that $c_n\in C^5(\overline{\Omega}_2)$, satisfy condition (30) and also

$c_n \in [1,d], c_n(x) = 1$ in $\Omega_2 \diagdown \Omega$. To make sure that $c_n \in C^5\left(\overline{\Omega}_2\right)$, we use the embedding theorem for $k = 2,3$. Hence, in our case the Hilbert space $H^{(1)} = H^7(\Omega_2) \subset C^5\left(\overline{\Omega}_2\right)$. The Hilbert space $H_1^{(1)}$ should be such that its norm would be stronger than the norm in $H^{(1)}$. Hence, we set $H_1^{(1)} := H^8(\Omega_2)$. Obviously $\overline{H}_1^{(1)} = H^{(1)}$ and any bounded set in $H_1^{(1)}$ is a precompact set in $H^{(1)}$. We define the set $G^{(1)} \subset H^{(1)}$ as

$$G^{(1)} = \left\{ c \in H^{(1)} : c \in [1,d], c(x) = 1 \text{ in } \Omega_2 \diagdown \Omega, \|\nabla c\|_{C(\overline{\Omega})} \leq A \text{ and (30) holds} \right\}, \quad (34)$$

where $A > 0$ is a constant such that

$$A \geq \frac{\overline{c} + 0.5}{\max_{\overline{\Omega}} |x|}.$$

Hence, the inequality $\|\nabla c\|_{C(\overline{\Omega})} \leq A$ does not contradict (30). Clearly, $G^{(1)}$ is a closed set which can be obtained as a closure of an open set.

To correctly introduce the operator F, we should carefully work with appropriate Hilbert spaces. This is because we should keep in mind the classical theory of hyperbolic initial boundary value problems, which guarantees smoothness of solutions of these problems [17]. Consider first the operator $\widehat{F} : G^{(1)} \to H^2(Q_T)$ defined as

$$\widehat{F}(c) = u_\varepsilon(x,t,c), (x,t) \in Q_T,$$

where $u_\varepsilon(x,t,c)$ is the solution of the problem (32), (33) with this specific coefficient c.

Lemma 2. *Let $T \geq T_1(\Omega, \overline{c}, d) > 0$, where T_1 was chosen in Theorems 3 and 4. Let $\Omega_2 = \Omega_2(T)$ be the domain chosen in Lemma 1 and $Q_T = \Omega_2 \times (0,T)$. Then the operator $\widehat{F}$ is Lipschitz continuous on $G^{(1)}$.*

Proof. Consider two functions $c_1, c_2 \in G^{(1)}$. Denote $\widetilde{c} = c_1 - c_2, \widetilde{u}(x,t) = u_\varepsilon(x,t,c_1) - u_\varepsilon(x,t,c_2)$. Then (28), (32) and (33) imply

$$\begin{aligned} c_1 \widetilde{u}_{tt} - \Delta \widetilde{u} &= -\widetilde{c} \partial_t^2 u_\varepsilon(x,t,c_2) \text{ in } Q_T, \\ \widetilde{u}(x,0) &= \widetilde{u}_t(x,0) = 0, \\ \widetilde{u} \mid_{\partial\Omega_2 \times (0,T)} &= 0. \end{aligned} \quad (35)$$

Let $f(x,t) = -\widetilde{c}\partial_t^2 u_\varepsilon(x,t,c_2)$ be the right hand side of equation (35). Since in fact $u_\varepsilon \in H^6(Q_T)$ (Lemma 1), then

$$\partial_t^k f \in L_2(Q_T), k = 0,...,4. \quad (36)$$

Hence, using Theorem 4.1 of Chapter 4 of [17], we obtain

$$\|\widetilde{u}\|_{H^2(Q_T)} \leq B_1 \|\widetilde{c}\|_{C(\overline{\Omega})} \left(\left\|\partial_t^2 u_\varepsilon(x,t,c_2)\right\|_{L_2(Q_T)} + \left\|\partial_t^3 u_\varepsilon(x,t,c_2)\right\|_{L_2(Q_T)} \right), \quad (37)$$

where the constant $B = B(A,d) > 0$ depends only on numbers A, d. Since By the embedding theorem $H^7(\Omega) \subset C^5\left(\overline{\Omega}\right)$ and $\|\cdot\|_{C^5(\overline{\Omega})} \leq C \|\cdot\|_{H^7(\Omega)}$ with a constant $C > 0$ depending only on Ω. Since $\widetilde{c} \in H^7(\Omega)$ and the norm $\|\widetilde{c}\|_{C(\overline{\Omega})}$ is weaker than the norm $\|\widetilde{c}\|_{C^5(\overline{\Omega})}$, then it is sufficient now to estimate norms $\|\partial_t^2 u_\varepsilon(x,t,c_2)\|_{L_2(Q_T)}, \|\partial_t^3 u_\varepsilon(x,t,c_2)\|_{L_2(Q_T)}$ from the above uniformly for all functions $c_2 \in G$, i.e. by a constant, $\|\partial_t^2 u_\varepsilon(x,t,c_2)\|_{L_2(Q_T)} + \|\partial_t^3 u_\varepsilon(x,t,c_2)\|_{L_2(Q_T)} \leq B_2$,where $B_2 = B_2\left(G^{(1)}\right) > 0$ is a constant depending only on the set $G^{(1)}$. And the latter estimate follows from Theorem 4.1 of Chapter 4 of [17] and the discussion after it. $\square$

Remark 2. Using (37) as well as the discussion presented after Theorem 4.1 of Chapter 4 of [17], one might try to estimate $\widetilde{u}$ in a stronger norm. However, since the function $\widetilde{c}\partial_t^3 u_\varepsilon(x,0,c_2) = \widetilde{c}\Delta f_\varepsilon(x-x_0) \neq 0$ in Ω, then this attempt would require a higher smoothness of functions $c_1, \widetilde{c}$. In any case, the estimate (37) is sufficient for our goal.

We define the operator $F^{(1)} : G^{(1)} \to L_2(\partial\Omega \times (0,T))$ as

$$F^{(1)}(c)(x,t) := g_\varepsilon = u_\varepsilon \mid_{\partial\Omega\times(0,T)} \in L_2(\partial\Omega \times (0,T)). \tag{38}$$

Lemma 3. *The operator $F^{(1)}$ is Lipschitz continuous and one-to-one on the set $G^{(1)}$.*

Proof. The Lipschitz continuity follows immediately from the trace theorem and Lemma 2. We prove now that $F^{(1)}$ is one-to-one. Suppose that $F^{(1)}(c_1) = F^{(1)}(c_2)$. Hence, $u_\varepsilon(x,t,c_1) \mid_{\partial\Omega\times(0,T)} = u_\varepsilon(x,t,c_2) \mid_{\partial\Omega\times(0,T)}$. Let $\widetilde{u}(x,t) = u_\varepsilon(x,t,c_1) - u_\varepsilon(x,t,c_2)$. Since $c_1 = c_2 = 1$ in $\mathbb{R}^k \diagdown \Omega$, then we have

$$\begin{aligned} \widetilde{u}_{tt} - \Delta\widetilde{u} &= 0 \text{ in } \left(\mathbb{R}^k \diagdown \Omega\right) \times (0,T), \\ \widetilde{u}(x,0) &= \widetilde{u}_t(x,0) = 0, \text{in } \mathbb{R}^k \diagdown \Omega, \\ \widetilde{u} &\mid_{\partial\Omega\times(0,T)} = 0. \end{aligned}$$

Hence, $\widetilde{u}(x,t) = 0$ in $\left(\mathbb{R}^k \diagdown \Omega\right) \times (0,T)$. Hence, we obtain

$$\widetilde{u} \mid_{\partial\Omega\times(0,T)} = \partial_n \widetilde{u} \mid_{\partial\Omega\times(0,T)} = 0. \tag{39}$$

Since by Lemma 1 both functions $u_\varepsilon(x,t,c_1), u_\varepsilon(x,t,c_2) \in C^3\left(\overline{\Omega} \times [0,T]\right)$, then (39) and Theorem 4 imply that $c_1(x) \equiv c_2(x)$. $\square$

Suppose that there exists the exact solution $c^* \in H_1^{(1)} \cap G^{(1)}$ of the equation

$$F^{(1)}(c^*) = g^*(x,t) := u_\varepsilon(x,t,c^*) \mid_{\partial\Omega\times(0,T)}. \tag{40}$$

Then for a given function $g^*(x,t)$ this solution is unique by Lemma 3. Let the function $\widetilde{g}(x,t) \in L_2(\partial\Omega \times (0,T))$ be such that

$$\widetilde{g}(x,t) = u_\varepsilon(x,t,c^*) \mid_{\partial\Omega\times(0,T)} + g_\delta(x,t), (x,t) \in \partial\Omega \times (0,T), \tag{41}$$

where the function $g_\delta(x,t)$ represents the error in the data,

$$\|g_\delta\|_{L_2(\partial\Omega\times(0,T))} \leq \delta. \tag{42}$$

Let the function $c_0 \in H_1^{(1)} \cap G^{(1)}$ be a first guess for c^*. We now introduce the regularization functional $J_\alpha^{(1)}$ as

$$J_\alpha^{(1)}(c) = \int_0^T \int_{\partial\Omega} \left(u_\varepsilon(x,t,c) - \widetilde{g}(x,t)\right)^2 dS_x dt + \alpha \left\|c - c_0\right\|_{H^8(\Omega)}^2 ; c, c_0 \in G, \tag{43}$$

where the dependence $\alpha = \alpha(\delta) = \delta^{2\mu}, \mu \in (0,1)$ is the same as in (9), also see Remark 1.

Let $m = m(\delta) = \inf_{H_1^{(1)} \cap G^{(1)}} J_\alpha^{(1)}(c)$ and $\{c_n\}_{n=1}^\infty \subset H_1^{(1)} \cap G^{(1)}$ be the corresponding minimizing sequence, $\lim_{n\to\infty} J_\alpha(c_n) = m(\delta)$. Similarly with (10) we introduce the set $K^{(1)}(\delta, x_0) \subset H_1^{(1)}$, which is a precompact set in $H^{(1)}$,

$$K^{(1)}(\delta, c_0) = \left\{ c \in G^{(1)} : \|c\|_1 \leq \sqrt{\delta^{2(1-\mu)} + \|c^* - c_0\|_1^2} + \|c_0\|_1 \right\}. \tag{44}$$

Let $\overline{K}^{(1)} = \overline{K}^{(1)}(\delta, c_0) \subset H^{(1)}$ be the closure of the set $K^{(1)}(\delta, c_0)$ in the norm of the space $H^{(1)}$. Hence, $\overline{K}^{(1)}(\delta, c_0)$ is a closed compact in the space $H^{(1)}$.

Theorem 6. *Let Hilbert spaces $H^{(1)}, H_1^{(1)}$ be ones defined above and the set $G^{(1)}$ be the one defined in (34). Let the number $T_1 = T_1(\Omega, \overline{c}, d) > 0$ be the one chosen in Theorems 3,4 and $T \geq T_1$. Assume that there exists the exact solution $c^* \in H_1^{(1)} \cap G^{(1)}$ of equation (40) and let $c_0 \in H_1^{(1)} \cap G^{(1)}$ be a first approximation for this solution. Suppose that (41) and (42) hold. Let $m = m(\delta) = \inf_{H_1^{(1)} \cap G^{(1)}} J_\alpha^{(1)}(c)$ and $\{c_n\}_{n=1}^\infty \subset H_1^{(1)} \cap G^{(1)}$ be the corresponding minimizing sequence, $\lim_{n\to\infty} J_{\alpha(\delta)}^{(1)}(c_n) = m(\delta)$, where the functional $J_{\alpha(\delta)}^{(1)}$ is defined in (43) and the dependence $\alpha = \alpha(\delta)$ is given in (9) (also, see Remark 1). Then for any number $\xi \in (0,1)$ there exists a sufficiently small number $\delta_0 = \delta_0(\xi) \in (0,1)$ such that for all $\delta \in (0, \delta_0)$ the following inequality holds*

$$\|c_n - c^*\|_{H^{(1)}} \leq \xi \|c_0 - c^*\|_{H_1^{(1)}}.$$

In addition, if a regularized solution $x_{\alpha(\delta)}$ exists for a certain $\delta \in (0, \delta_0)$, i.e. $J_{\alpha(\delta)}\left(c_{\alpha(\delta)}\right) = m(\delta)$, then

$$\left\|c_{\alpha(\delta)} - c^*\right\|_{H^{(1)}} \leq \xi \|c^* - c_0\|_{H_1^{(1)}}.$$

Proof. Let $\overline{K}^{(1)}$ be the above closed compact. It follows from Theorem 1 and Lemma 3 that the operator $\left(F^{(1)}\right)^{-1} : F^{(1)}\left(\overline{K}^{(1)}\right) \to \overline{K}^{(1)}$ exists and is continuous. Let $\omega_{F^{(1)}}$ be its modulus of the continuity. Then similarly with the proof of Theorem 2 we obtain

$$\|c_n - c^*\|_{H^{(1)}} \leq \omega_{F^{(1)}}\left(\sqrt{\delta^2 + \delta^{2\mu}\|c^* - c_0\|_{H_1^{(1)}}^2} + \delta\right).$$

The rest of the proof repeats the corresponding part of the proof of Theorem 2. □

Remarks 3:

1. When applying the technique of [7,11-15] for proofs of uniqueness theorems for CIPs, like, e.g. Theorem 3, it is often possible to obtain Hölder or even Lipschitz estimates for the modulus of the continuity of the corresponding operator. The Hölder estimate means $\omega_F(z) \leq Cz^\nu, \nu \in (0,1)$ and the Lipschitz estimate means $\omega_F(z) \leq Cz$ with a constant $C > 0$. This is because Carleman estimates imply at least Hölder estimates, see Chapter 2 in [14]. This thought was used in [9] for a purpose, which is different from the one of this paper (see some comments in Introduction). In particular, the Lipschitz stability for a similar inverse problem for the equation $u_{tt} = div\,(a(x)\nabla u)$ was obtained in [15]. So, in the case of, e.g. the Lipschitz stability and under a reasonable assumption that $\delta \leq \delta^\mu \|c^* - c_0\|^2_{H_1^{(1)}}$ (see (4)) the last estimate of the proof of Theorem 6 would become

$$\|c_n - c^*\|_{H^{(1)}} \leq C\delta^\mu \|c^* - c_0\|^2_{H_1^{(1)}}.$$

Since it is usually assumed that $\delta^\mu << 1$, then this is an additional indication of the fact that the accuracy of a regularized solution is significantly better than the accuracy of the first guess c_0 even for a *single* value of the regularization parameter, as opposed to the limiting case of $\delta, \alpha(\delta) \to 0$ of the classical regularization theory. The only reason why we do not provide such detailed estimates here is that it would be quite space consuming to do so.

2. A conclusion, which is very similar to the above first remark, can be drawn for the less restrictive finite dimensional case considered in subsection 3.3. Since all norms in a finite dimensional space are equivalent, then in this case norms $\|c_n - c^*\|_{H^{(1)}}, \|c^* - c_0\|_{H_1^{(1)}}$ can be replaced with $\|c_n - c^*\|_{L_2(\Omega)}, \|c^* - c_0\|_{L_2(\Omega)}$, which is more convenient, see Theorem 7 in subsection 3.3.

3. The smoothness requirement $c_n \in H^8(\Omega_2)$ is an inconvenient one. However, given the absence of proper uniqueness results for the original Inverse Problem 1, there is nothing what can be done at this point. Also, this smoothness requirement is imposed only to make sure that Lemma 1 is valid and we need the smoothness of Lemma 1 for Theorem 3. In the next subsection we present a scenario, which is more realistic from the computational standpoint. However, instead of relying on the rigorously established (in [13,14]) uniqueness Theorem 3, we only assume in subsection 3.3 that the uniqueness holds.

3.3 The finite dimensional case

Because of a number of inconveniences of the infinitely dimensional case mentioned in subsection 2.2 and in Remark 3, we now consider the finite dimensional case, which is more realistic for computations. Fix a certain sufficiently large number $T > 0$ and let $\Omega_2 = \Omega_2(T)$ be the domain chosen in Lemma 1. Since in the works [2-6,16] and in the numerical section 4 only standard triangular or tetrahedral finite elements are applied and these elements form a finite dimensional subspace of piecewise linear functions in the space $L_2(\Omega)$, we assume in this subsection that

$$c \in C\left(\overline{\Omega}_2\right) \cap H^1(\Omega_2) \text{ and derivatives } c_{x_i} \text{ are bounded in } \overline{\Omega}_2, i = 1, ..., k, \quad (45)$$

$$c(x) \in [1, d] \text{ in } \Omega, c(x) = 1 \text{ in } \Omega_2 \diagdown \Omega. \quad (46)$$

Let $H^{(2)} \subset L_2(\Omega_2)$ be a finite dimensional subspace of the space $L_2(\Omega_2)$ in which all functions satisfy conditions (45). Define the set $G^{(2)}$ as

$$G^{(2)} = \left\{ c \in H^{(2)} \text{ satisfying } (46) \right\}. \tag{47}$$

Since norms $\|\cdot\|_{C(\overline{\Omega})}$ and $\|\cdot\|_{L_2(\Omega)}$ are equivalent in the finite dimensional space $H^{(2)}$ and $G^{(2)}$ is a closed bounded set in $C\left(\overline{\Omega}\right)$, then $G^{(2)}$ is also a closed set in $H^{(2)}$.

Let $\theta \in (0,1)$ be a sufficiently small number. We follow [4,6] by replacing the function $\delta(x-x_0)$ in (21) with the function $\delta_\theta(x-x_0)$,

$$\delta_\theta(x-x_0) = \begin{cases} C_\theta \exp\left(\frac{1}{|x-x_0|^2-\theta^2}\right), & |x-x_0| < \theta, \\ 0, & |x-x_0| \geq \theta, \end{cases}$$

where the constant C_θ is such that

$$\int_{\mathbb{R}^k} \delta_\theta(x-x_0)\,dx = 1.$$

We assume that θ is so small that

$$\delta_\theta(x-x_0) = 0 \text{ for } x \in \Omega \cup \left(\mathbb{R}^k \diagdown \Omega_1\right). \tag{48}$$

Let $T>0$ be a certain number. For any function $c \in G^{(2)}$ consider the solution $u_\theta(x,t)$ of the following Cauchy problem

$$c(x)\,\partial_t^2 u_\theta = \Delta u_\theta \text{ in } \mathbb{R}^k \times (0,T), \tag{49}$$

$$u_\theta(x,0) = 0, \partial_t u_\theta(x,0) = \delta_\theta(x-x_0). \tag{50}$$

Similarly with the Cauchy problem (20), (27) we conclude that because of (48), there exists unique solution of the problem (49), (50) such that $u_\theta \in H^2(\Phi \times (0,T))$ for any bounded domain $\Phi \subset \mathbb{R}^k$. In addition, $u_\theta(x,t) = 0$ in $\left(\mathbb{R}^k \diagdown \Omega_2\right) \times (0,T)$.

Similarly with (38) we introduce the operator $F^{(2)} : G^{(2)} \to L_2(\partial\Omega \times (0,T))$ as

$$F^{(2)}(c)(x,t) := g_\theta = u_\theta\mid_{\partial\Omega\times(0,T_3)} \in L_2(\partial\Omega \times (0,T)). \tag{51}$$

Analogously to Lemma 3 one can prove

Lemma 4 (see Lemma 7.1 of [6] for an analogous result). *The operator $F^{(2)}$ is Lipschitz continuous.*

Unlike the previous two subsections, there is no uniqueness result for the operator $F^{(2)}$. The main reason of this is that $\delta_\theta(x-x_0) = 0$ in Ω, unlike the function $f_\varepsilon(x-x_0)$, see Introduction for the discussion of uniqueness results. In addition the function $c(x)$ is not sufficiently smooth now. Hence, we simply impose the following

Assumption. *The operator $F^{(2)} : G^{(2)} \to L_2(\partial\Omega \times (0,T))$ is one-to-one.*

Similarly with (40)-(42) we assume that there exists the exact solution $c^* \in G^{(2)}$ of the equation

$$F^{(2)}(c^*) = g^*(x,t) := u_\theta(x,t,c^*) \mid_{\partial\Omega\times(0,T)} . \tag{52}$$

It follows from assumption that, for a given function $g^*(x,t)$, this solution is unique. Let the function $\overline{g}(x,t) \in L_2(\partial\Omega \times (0,T))$. We assume that

$$\overline{g}(x,t) = u_\theta(x,t,c^*) \mid_{\partial\Omega\times(0,T)} + g_\delta(x,t), (x,t) \in \partial\Omega \times (0,T), \tag{53}$$

where the function $g_\delta(x,t)$ represents the error in the data,

$$\|g_\delta\|_{L_2(\partial\Omega\times(0,T))} \le \delta. \tag{54}$$

Let $\alpha = \alpha(\delta) = \delta^{2\mu}, \mu \in (0,1)$ be the same as in (9), also see Remark 1. We now introduce the regularization functional $J_\alpha^{(2)}$ as

$$J_\alpha^{(2)}(c) = \int_0^T \int_{\partial\Omega} (u_\theta(x,t,c) - \overline{g}(x,t))^2 \, dS_x dt + \alpha \|c - c_0\|_{L_2(\Omega)}^2 ; \; c, c_0 \in G^{(2)}. \tag{55}$$

Obviously this functional is more convenient to work with than the functional $J_\alpha^{(1)}(c)$ in (43). This is because the $H^8(\Omega)$ norm of (43) is replaced with the $L_2(\Omega)$ norm in (55). However, unlike the previous subsection where the one-to-one property of the operator $F^{(1)}$ was rigorously guaranteed, here we only assume this property for the operator $F^{(2)}$.

Theorem 7. *Let $H^{(2)} \subset L_2(\Omega_2)$ be the above finite dimensional subspace of the space $L_2(\Omega_2)$, in which all functions satisfy conditions (45). Let the set $G^{(2)} \subset H^{(2)}$ be the one defined in (47). Let the function u_θ be the solution of the Cauchy problem (49), (50) such that $u_\theta \in H^2(\Phi \times (0,T))$ for any bounded domain $\Phi \subset R^k$. Let $F^{(2)}$ be the operator defined in (51) and let Assumption be true. Assume that there exists the exact solution $c^* \in G^{(2)}$ of equation (52) and let $c_0 \in G^{(2)}$ be a first guess for this solution. Suppose that (53) and (54) hold. Let $m = m(\delta) = \min_{\overline{G}^{(2)}} J_\alpha^{(2)}(c)$. Then a regularized solution $c_{\alpha(\delta)} \in G^{(2)}$ exists, i.e.*

$$m = m(\delta) = \min_{G^{(2)}} J_\alpha^{(2)}(c) = J_\alpha^{(2)}\left(c_{\alpha(\delta)}\right). \tag{56}$$

Also, for any number $\xi \in (0,1)$ there exists a sufficiently small number $\delta_0 = \delta_0(\xi) \in (0,1)$ such that for all $\delta \in (0,\delta_0)$ the following inequality holds

$$\left\|c_{\alpha(\delta)} - c^*\right\|_{L_2(\Omega)} \le \xi \|c^* - c_0\|_{L_2(\Omega)}.$$

Proof. Since $G^{(2)} \subset H^{(2)}$ is a closed bounded set then the existence of a minimizer $c_{\alpha(\delta)}$ of the functional $J_\alpha^{(2)}(c)$ is obvious even for any value of $\alpha > 0$. The set $G^{(2)} \subset H^{(2)}$ is a closed compact. Hence, it follows from Theorem 1 and Assumption that the operator

$\left(F^{(2)}\right)^{-1}: F^{(2)}\left(G^{(2)}\right) \rightarrow G^{(2)}$ exists and is continuous. Let $\omega_{F^{(2)}}$ be its modulus of the continuity. Then similarly with the proof of Theorem 2 we obtain

$$\left\|c_{\alpha(\delta)}-c^{*}\right\|_{L_2(\Omega)} \leq \omega_{F^{(2)}}\left(\sqrt{\delta^{2}+\delta^{2\mu}\left\|c^{*}-c_0\right\|_{L_2(\Omega)}^{2}}+\delta\right).$$

The rest of the proof is the same as the corresponding part of the proof of Theorem 2. □

Although the existence of a regularized solution is established in Theorem 7, it is still unclear how many such solutions exist and how to obtain them in practical computations. We show now that if the function c_0 is sufficiently close to the exact solution c^*, then the regularized solution is unique and can be found by a version of the gradient method without a risk of coming up with a local minima or a ravine. In doing so we use results of [6].

Let $\mu_1 \in (0,1)$ be such a number that

$$2\mu \in \left(0, \min\left(\mu_1, 2\left(1-\mu_1\right)\right)\right), \tag{57}$$

where the number μ was defined in (9). Assume that

$$\left\|c_0-c^{*}\right\|_{L_2(\Omega)} \leq \delta^{\mu_1}. \tag{58}$$

Let the number $\beta \in (0,1)$ be independent on δ. Denote

$$\begin{aligned} V_{\left(1+\sqrt{2}\right)\delta^{\mu_1}}\left(c^{*}\right) &= \left\{c \in G^{(2)}: \left\|c-c^{*}\right\|_{L_2(\Omega)} < \left(1+\sqrt{2}\right)\delta^{\mu_1}\right\}, \\ V_{\beta\delta^{2\mu}}\left(c_{\alpha(\delta)}\right) &= \left\{c \in G^{(2)}: \left\|c-c_{\alpha(\delta)}\right\|_{L_2(\Omega)} < \beta\delta^{2\mu}\right\}. \end{aligned}$$

Note that it follows from (57) that $\beta\delta^{2\mu} >> \delta^{\mu_1}$ for sufficiently small δ, since

$\lim_{\delta\rightarrow 0}\left(\beta\delta^{2\mu}/\delta^{\mu_1}\right)=\infty$. The following theorem can be derived from a simple reformulation of Theorem 7.2 of [6] for our case.

Theorem 8. *Assume that conditions of Theorem 7 hold. Also, let (57) and (58) be true. Let $c_{\alpha(\delta)}$ be a regularized solution defined in (56). Then there exists a sufficiently small number $\delta_1 \in (0,1)$ and a number $\beta \in (0,1)$ independent on δ such that for any $\delta \in (0,\delta_1]$ the set $V_{\left(1+\sqrt{2}\right)\delta^{\mu_1}}\left(c^{*}\right) \subset V_{\beta\delta^{2\mu}}\left(c_{\alpha(\delta)}\right)$, the functional $J_{\alpha(\delta)}^{(2)}(c)$ is strongly convex on the set $V_{\beta\delta^{2\mu}}\left(c_{\alpha(\delta)}\right)$ and has the unique minimizer $\overline{c}_{\alpha(\delta)}$ on the set $V_{\left(1+\sqrt{2}\right)\delta^{\mu_1}}\left(c^{*}\right)$. Furthermore, $\overline{c}_{\alpha(\delta)}=c_{\alpha(\delta)}$. Hence, since by (58) $c_0 \in V_{\left(1+\sqrt{2}\right)\delta^{\mu_1}}\left(c^{*}\right)$, then any gradient-like method of the minimization of the functional $J_{\alpha(\delta)}^{(2)}(c)$, which starts from the first guess c_0, converges to the regularized solution $c_{\alpha(\delta)}$.*

Remarks 4:

1. Thus, Theorems 7 and 8 emphasize once again the importance of obtaining a good first approximation for the exact solution, i.e. the importance of globally convergent numerical methods. Indeed, if such a good approximation is available, Theorem 8 guarantees convergence of any gradient-like method to the regularized solution and Theorem 7 guarantees that this solution is more accurate than the first approximation.

2. As it was mentioned in Introduction, global convergence theorems of [2,3] in combination with Theorems 7 and 8 guarantee that the two-stage numerical procedure of [2-6] converges globally. Another confirmation of this can be found in first and second Remarks 3.

4 Numerical Example

We now briefly describe a numerical example for Inverse Problem 1. The sole purpose of this section is to illustrate how Theorems 7 and 8 work. Figures 1-a), 1-b) and 1-c) of this example are published in [4], and we refer to this reference for details. So, first we apply the globally convergent numerical method of [2,3]. As a result, we obtain a good first approximation for the solution of Inverse Problem 1. Next, we apply the locally convergent adaptivity technique for refinement. In doing so, we use the solution obtained on the first stage as the starting point. In this example we use the incident plane wave rather than the point source. This is because the case of the plane wave works better numerically than the point source. In fact, we have used the point source in our theory both here and in [2-6] only to obtain a certain asymptotic behavior of the Laplace transform of the function $u(x,t)$, see Lemma 2.1 in [2]. We need this behavior for some theoretical results. However, in the case of the plane wave we verify this asymptotic behavior computationally, see subsection 7.2 in [2]. As it was stated above, numerical studies were conducted using the standard triangular finite elements when solving Inverse Problem 1.

The big square on Figure 1-a) depicts the domain Ω. Two small squares display two inclusions to be imaged. In this case we have

$$c(x)=\left\{\begin{array}{c} 4 \text{ in small squares,} \\ 1 \text{ outside of small squares.}\end{array}\right. \tag{59}$$

The plane wave falls from the top. The scattered wave $g(x,t)=u\mid_{\partial\Omega\times(0,T)}$ in (22) is known at the boundary of the big square. The multiplicative random noise of the 5% level was introduced in the function $g(x,t)$. Hence, in our case $\delta\approx 0.05$. We have used $\alpha=0.01$ in (55). This value of the regularization parameter was chosen by trial and error. Our algorithm does not use a knowledge of background values of the function $c(x)$, i.e. it does not use a knowledge of $c(x)$ outside of small squares.

Figure 1-b) displays the result of the performance of the first stage of our two-stage numerical procedure. The computed function $c_{glob}(x):=c_0(x)$ has the following values: the maximal value of $c_0(x)$ within each of two imaged inclusions is 3.8. Hence, we have only 5% error (4/3.8) in the imaged inclusion/background contrast, see (59). Also, $c_0(x)=1$ outside of these imaged inclusions. Figure 1-c) displays the final image. It is obvious that the image of Figure 1-b) is refined, just as it was predicted by Theorem 7. Indeed, locations of both imaged inclusions are accurate. Let $c_\alpha(x)$ be the computed coefficient. Its maximal value is $\max c_\alpha(x)=4$ and it is achieved within both imaged inclusions. Also, $c_\alpha(x)=1$ outside of imaged inclusions.

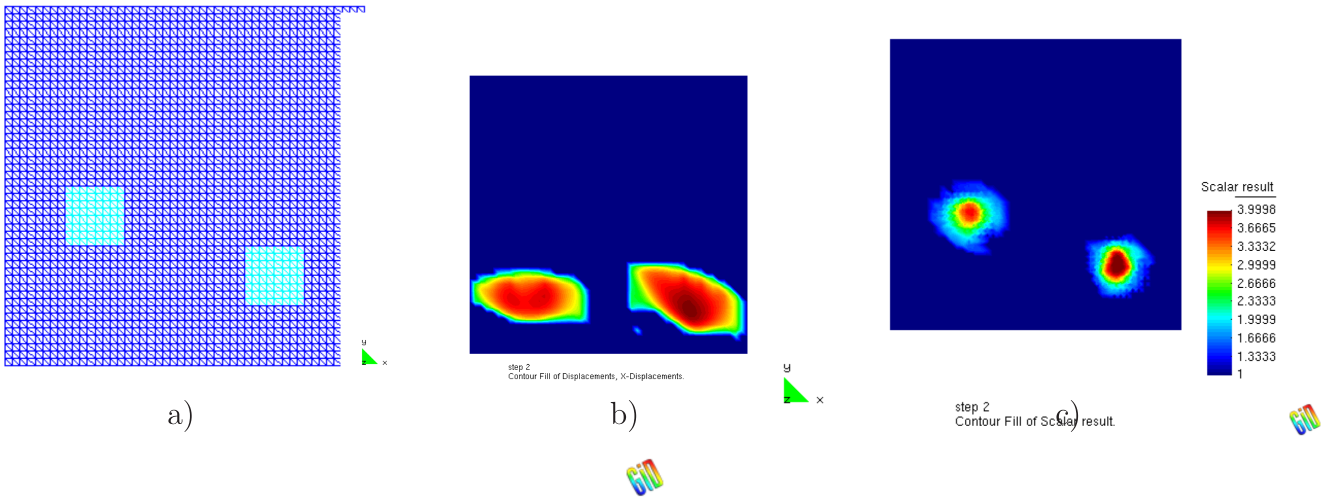

Figure 1: See details in the text of section 4. a) The big square depicts the domain Ω. Two small squares display two inclusions to be imaged, see (59) for values of the function $c(x)$. The plane falls from the top. The scattered wave is known at the boundary of the big square. b) The result obtained by the globally convergent first stage of the two-stage numerical procedure of []. c) The result obtained after applying the second stage. A very good refinement is achieved, just as predicted by Theorem 7. Both locations of two inclusions and values of the function $c(x)$ inside and outside of them are imaged with a very good accuracy.

Acknowledgments

This work was supported by: (1) The US Army Research Laboratory and US Army Research Office grant W911NF-08-1-0470, (2) the grant RFBR 09-01-00273a from the Russian Foundation for Basic Research, (3) the Swedish Foundation for Strategic Research (SSF) at the Gothenburg Mathematical Modeling Center (GMMC), and (4) the Visby Program of the Swedish Institute.

References

[1] 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, 2001.

[2] L. Beilina and M.V. Klibanov, A globally convergent numerical method for a coefficient inverse problem, *SIAM J. Sci. Comp.*, 31, 478-509, 2008.

[3] 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.

[4] 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.

[5] L. Beilina and M.V. Klibanov, Reconstruction of dielectrics from experimental data via a hybrid globally convergent/adaptive inverse algorithm, submitted for publication, a preprint is available on-line at http://www.ma.utexas.edu/mp_arc/

[6] 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.

[7] A.L. Buhgeim and M.V. Klibanov, Uniqueness in the large of a class of multidimensional inverse problems, *Soviet Math. Doklady*, 17, 244-247, 1981

[8] M. Cheney and D. Isaacson, Inverse problems for a perturbed dissipative half-space, *Inverse Problems,* 11, 865-888, 1995.

[9] J. Cheng and M. Yamamoto, One new strategy for a priori choice of regularizing parameters in Tikhonov's regularization, *Inverse Problems*, 16, L31-L38, 2000.

[10] H.W. Engl, M. Hanke and A. Neubauer, *Regularization of Inverse Problems*, Kluwer Academic Publishers, Boston, 2000.

[11] M.V. Klibanov, Uniqueness of solutions in the 'large' of some multidimensional inverse problems. In *Non-Classical Problems of Mathematical Physics*, Proc. of Computing Center of the Siberian Branch of USSR Academy of Science, Novosibirsk, 101-114, 1981.

[12] M.V. Klibanov, Inverse problems in the 'large' and Carleman bounds, *Differential Equations*, 20, 755-760, 1984.

[13] M.V. Klibanov, Inverse problems and Carleman estimates, *Inverse Problems,* 8, 575–596, 1991.

[14] M.V. Klibanov and A. Timonov, *Carleman Estimates for Coefficient Inverse Problems and Numerical Applications*, VSP, Utrecht, The Netherlands, 2004.

[15] M.V. Klibanov and M. Yamamoto, Lipschitz stability of an inverse problem for an acoustic equation, *Applicable Analysis*, 85, 515-538, 2006.

[16] 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.

[17] O.A. Ladyzhenskaya, *Boundary Value Problems of Mathematical Physics*, Springer, Berlin, 1985.

[18] A.N. Tikhonov, A.V. Goncharsky, V.V. Stepanov and A.G. Yagola, 1995 *Numerical Methods for the Solution of Ill-Posed Problems*, Kluwer, London, 1995.

[19] M. Yamamoto, Carleman estimates for parabolic equations and applications, *Inverse Problems,* 25, 123013, 2009.