NORGES TEKNISK-NATURVITENSKAPELIGE
UNIVERSITET

A globally convergent numerical method and adaptivity for a hyperbolic coefficient inverse problem

by

Larisa Beilina$^{\triangledown}$, Michael V. Klibanov*

PREPRINT
NUMERICS NO. 1/2009

NORWEGIAN UNIVERSITY OF
SCIENCE AND TECHNOLOGY
TRONDHEIM, NORWAY

This report has URL
http://www.math.ntnu.no/preprint/numerics/2009/N1-2009.pdf
Address: Department of Mathematical Sciences, Norwegian University of Science and Technology, N-7491 Trondheim, Norway.

A globally convergent numerical method and adaptivity for a hyperbolic coefficient inverse problem

Larisa Beilina[▽], Michael V. Klibanov[*]

February 4, 2009

A globally convergent numerical method for a multidimensional Coefficient Inverse Problem for a hyperbolic equation is presented. It is shown that this technique provides a good starting point for the so-called finite element adaptive method (adaptivity). This leads to a natural two-stage numerical procedure, which synthesizes both these methods. Numerical examples are presented.

1 Introduction

This paper is a continuation of the previous publication of the authors [5], where a new globally convergent numerical method for a Coefficient Inverse Problem (CIP) for a hyperbolic PDE was developed. Compared with [5], the main new element here is a synthesis of the method of [5] with the locally convergent so-called finite element adaptive method, which we call "adaptivity" for brevity. The adaptivity technique for inverse problems was previously developed in [6, 7, 8, 9, 4]. The underlying reason of this synthesis is that the estimate of the difference between the correct solution and the computed one in the global convergence Theorem 6.1 depends on a small positive parameter η. This parameter incorporates both the error in the boundary data and errors generated by some approximations of the numerical procedure of [5]. The error in the boundary data models the error in measurements and is, therefore unavoidable. At the same time, two other approximation errors cannot be made zero, and they are not parts of previously developed locally convergent techniques. On the other hand, since η is small, then Theorem 6.1 guarantees that the solution obtained by the technique of [5] provides a good approximation for the correct solution of the CIP. Therefore, this solution can be used as a good starting point for a subsequent enhancement via a locally convergent numerical method, which is the adaptivity in our case. As a result, a natural two-stage numerical procedure is developed here. On the first stage, the globally convergent method of [5] provides a good approximation for the correct solution. And on the second stage, this approximation is taken as the starting point for the adaptivity technique, which provides an enhancement, i.e., a better approximation for the correct solution. The adaptivity technique is based on several applications

[*▽] Larisa Beilina,Department of Mathematical Sciences, Chalmers University of Technology and Gothenburg University, SE-42196 Gothenburg, Sweden, *email*: Larisa.Beilina@chalmers.se.

[*] Michael V. Klibanov, Department of Mathematics and Statistics University of North Carolina at Charlotte, Charlotte, NC 28223, USA, *email*: mklibanv@uncc.edu

of the quasi-Newton method. Convergence of Newton-like methods for general ill-posed problems was proven in [3].

We call a numerical method for a CIP *globally convergent* if: (1) a theorem is proven, which ensures that this method leads to a good approximation for the correct solution of that CIP, regardless on the availability of a priori given good first guess for that solution, and (2) this theorem is confirmed by numerical experiments. On the other hand, convergence of a *locally convergent* numerical method to the correct solution can be guaranteed only if the starting point is located in a small neighborhood of this solution.

There are four more new elements of this paper compared with [5]: (1) The globally convergent algorithm is different from one in [5] in the sense that now "inner" iterations with respect to terms in certain quasilinear elliptic equations are used until they converge. Whereas previously a priori chosen number of iterations was used. This change requires a modification of the proof of the convergence Theorem 6.1 compared with [5]. (2) The stopping rule for the globally convergent part differs from one of [5]. Namely, we now evaluate certain L_2 norms at the boundary rather than inside of the domain of interest. (3) The first rigorous explanation is presented of the meaning of the so-called ψ function in the adaptivity technique in an estimate of the difference between correct and computed solutions. This is unlike the previous heuristic argument of [9], where the function ψ was introduced for a CIP. (4) 2-D numerical examples are different from ones of [5].

A conventional way to solve numerically a CIP for a PDE is via the minimization of a least squares objective functional. This functional characterizes the misfit between the data and the solution of that PDE with a "guess" for the unknown coefficient. However, it is well known that the phenomenon of multiple local minima and ravines of these functionals represents the major obstacle in this approach. Because of this phenomenon, any gradient-like method of the minimization of such a functional would likely converge to a local minimum, which is located far from the correct solution. Furthermore, due to the ill-posed nature of CIPs, a global minimum, even a well pronounced one, is not necessarily close to the correct solution. Hence, there is no guarantee that the calculated coefficient is indeed close to the correct one. Hence, one needs to know a priori a good first approximation for the solution. At the same time, in many important applications such an approximation is unavailable. The method of [5] relies on the structure of the underlying differential operator instead of using a least squares objective functional. Thus, the phenomenon of local minima is avoided in this method.

The adaptivity technique minimizes least squares objective functionals on a sequence of adaptively refined meshes until images are stabilized (usually on 4-5 refined meshes). The minimization is performed via the quasi-Newton method. The key idea of the adaptivity is that on each step a posteriori analysis shows subdomains where the biggest error in the solution is. These are those subdomains where the gradient of the Lagrangian attains its maximal values (within certain range). An important point here is that those subdomains are found without a priori knowledge of the solution. Thus, additional finite elements are used in such subdomains. It was shown in previous publications that the adaptivity is capable to significantly improve reconstruction results. At the same time, it was shown numerically in the recent publication [10] that the adaptivity cannot provide good quality images unless a good first guess about the solution is known a priori. This is because the quasi-Newton method is a locally convergent one. The latter leads to a logical conclusion that a synthesis of the adaptivity with the globally convergent method of [5] should be used. In our numerical experiments we image a medium with small inclusions in it, although we do not assume a priori knowledge of such a structure. We refer to [1] and references cited there for another approach to imaging of small inclusions.

There are also some other numerical methods for multidimensional CIPs, which do not use a good first guess for the solution. While the current paper works with a single measurement event, they work for some CIPs with the data resulting from multiple measurements [12, 13, 14, 24, 25, 26]. These publications were discussed in [5].

In section 2 we formulate both forward and inverse problems. In section 3 we transform the inverse problem to the Dirichlet boundary value problem for a nonlinear integral differential equation in which the unknown coefficient is not present. Since this transformation was described in several previous publications [19, 5], we outline it only briefly here for the sake of completeness. It is the numerical solution of the resulting equation, which represents the major difficulty. The numerical method of this solution was the main new point of [5], In section 4 we formulate the layer stripping procedure with respect to $s > 0$, which is the parameter of the Laplace transform of the original hyperbolic PDE. Note that we do not use the inverse Laplace transform, since approximations for the unknown coefficient are obtained in the "Laplace's domain". In section 5 we describe the algorithm. Section 6 is devoted to the convergence analysis. In section 7 we briefly describe the method of the solution of the forward problem. In section 8 we describe the main ideas of the adaptivity technique referring to the proof of a posteriori estimate to [7, 8, 9]. In particular, the above mentioned rigorous explanation of the meaning of the function ψ is given in subsection 8.2. In section 9 numerical experiments are presented. We summarize our results in section 10. Some procedures are outlined only briefly in this paper, since they were discussed in details in [5].

2 Statements of Forward and Inverse Problems

As the forward problem, we consider the Cauchy problem for a hyperbolic PDE. The case of a boundary value problem in a finite domain is not considered in our theoretical derivations only because an analogue of the asymptotic behavior (9) is not proved in this case, since (9) is actually derived from Theorem 4.1 of [29]. That theorem establishes a certain asymptotic behavior of the fundamental solution of a hyperbolic equation near the characteristic cone.

Consider the Cauchy problem for the hyperbolic equation

$$c(x)\, u_{tt} = \Delta u \text{ in } \mathbb{R}^3 \times (0, \infty), \tag{1}$$

$$u(x, 0) = 0, u_t(x, 0) = \delta(x - x_0). \tag{2}$$

Equation (1) governs a wide range of applications, including e.g., propagation of 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 in a non-magnetic medium, the dimensionless coefficient $c(x) = \varepsilon_r(x)$, where $\varepsilon_r(x)$ is the relative dielectric function of the medium, see [15], where this equation was derived from Maxwell's equations in the 2-D case. Let d_1 and d_2 be two positive constants and $\Omega \subset \mathbb{R}^3$ be a convex bounded domain with the boundary $\partial\Omega \in C^3$. We assume that the coefficient $c(x)$ of equation (1) is such that

$$c(x) \in [d_1, 2d_2], d_1 < d_2, c(x) = 2d_1 \text{ for } x \in \mathbb{R}^3 \setminus \Omega, \tag{3}$$

$$c(x) \in C^2\left(\mathbb{R}^3\right), \tag{4}$$

We consider the following

Inverse Problem. Suppose that the coefficient $c(x)$ satisfies (3) and (4), where the positive numbers d_1 and d_2 are 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 $x_0 \notin \overline{\Omega}$

$$u(x,t) = g(x,t), \forall (x,t) \in \partial\Omega \times (0,\infty). \tag{5}$$

A priori knowledge of constants d_1, d_2 corresponds well with the Tikhonov concept for ill-posed problems [30]. In applications the assumption $c(x) = 2d_1$ for $x \in \mathbb{R}^3 \setminus \Omega$ means that the target coefficient $c(x)$ has a known constant value outside of the medium of interest Ω. Another argument here is that one should bound the coefficient $c(x)$ from the below by a positive number to ensure that the operator in (1) is a hyperbolic one on all iterations of our method. Since we do not impose any "smallness" conditions on numbers d_1 and d_2, our numerical method is not a locally convergent one. The function $g(x,t)$ models time dependent measurements of the wave field at the boundary of the domain of interest. In practice measurements are performed at a number of detectors, of course. In this case the function $g(x,t)$ can be obtained via one of standard interpolation procedures, a discussion of which is outside the scope of this publication. In the case of a finite time interval, on which measurements are performed, one should assume that this interval is large enough and thus, the t-integral of the Laplace transform over this interval is approximately the same as one over $(0,\infty)$.

Consider the Laplace transform of the functions u,

$$w(x,s) = \int_0^\infty u(x,t)e^{-st}dt, \text{ for } s > \underline{s} = const. > 0, \tag{6}$$

where $\underline{s}$ is a certain number. It is sufficient to choose $\underline{s}$ such that the integral (6) would converge together with corresponding (x,t)-derivatives. We call the parameter s *pseudo frequency*. The equation for the function w is

$$\Delta w - s^2 c(x) w = -\delta(x - x_0) c(x_0), \forall s \geq \underline{s} = const. > 0 \tag{7}$$

with the following condition at the infinity

$$\lim_{|x|\to\infty} w(x,s) = 0, \forall s \geq \underline{s} = const. > 0. \tag{8}$$

Under some natural conditions linked with the regularity of geodesic lines generated by the eikonal equation corresponding to the function $c(x)$ the following asymptotic behavior takes place (see Lemma 2.1 in [5])

$$D_x^\beta D_s^\gamma w(x,s) = D_x^\beta D_s^\gamma \left\{ \frac{\exp[-sl(x,x_0)]}{f(x,x_0)} \left[1 + O\left(\frac{1}{s}\right) \right] \right\}, s \to \infty, \tag{9}$$

where $|\beta| \leq 2, \gamma = 0, 1, x \neq x_0, f(x,x_0)$ is a certain function, $f(x,x_0) \neq 0$ for $x \neq x_0$ and $l(x,x_0)$ is the length of the geodesic line connecting points x and x_0.

We briefly mention now that the idea of [5] can also be extended to similar CIPs for the parabolic PDE

$$\begin{aligned} c(x)U_t &= \Delta U - a(x)U, \\ U(x,0) &= \delta(x - x_0). \end{aligned} \tag{10}$$

To do this, one needs to apply the following analogue of the above Laplace transform

$$W(x,s) = \int_0^\infty U(x,t) \exp\left(-s^2 t\right) dt.$$

Hence, $\Delta W - \left(s^2 c(x) + a(x)\right) W = -\delta(x - x_0)$ and also W satisfies (8). In the electromagnetic case equation (10) governs propagation of a component of the electric field in a conductive medium with the conductivity function $\sigma(x) := c(x)$. In the case of diffuse optical tomography one can usually assume that the diffusion coefficient $D := 1/c \equiv const. > 0$ and the target of the CIP is the spatially changing absorption coefficient $\mu_a(x) := a(x)$, see, e.g., [2].

Although we have only one condition (5) rather than "traditional" two boundary conditions for our inverse problem, the information about the normal derivative of the function w at $\partial\Omega$ is actually inscribed in (5), because the original equation (6) holds in a wider domain and the coefficient $c(x)$ is known outside of Ω. To formalize the latter, one should consider the boundary value problem for equation (7) for $x \in \mathbb{R}^3 \setminus \Omega$ with the boundary condition (5) and condition (8). Solution of this problem provides the normal derivative of the function w at $\partial\Omega$. The question of uniqueness of this Inverse Problem is a well known long standing open problem. It is addressed positively only if the function $\delta(x - x_0)$ above is replaced with a such a function $f(x) \in C^\infty\left(\mathbb{R}^3\right)$ that $f(x) \neq 0, \forall x \in \overline{\Omega}$. An example of this function is function is

$$f(x) = C_\varepsilon \exp\left(-\frac{|x - x_0|^2}{\varepsilon^2}\right), \int\limits_{\mathbb{R}^3} f(x)\, dx = 1,$$

where $\varepsilon > 0$ is a small positive number and the positive constant C_ε is chosen such that the above integral equals 1. Corresponding theorems are proved via the method of Carleman estimates [19, 20]. In principle, one can replace the $\delta(x - x_0) -$ function with a $\delta(x - x_0) -$ like smooth function, which is not zero in $\overline{\Omega}$. The resulting function $\widetilde{w}$ will be close to the function w in a certain sense, and the above mentioned uniqueness result would be applicable then. In principle our numerical method can be extended to this case, although a corresponding development is outside of the scope of this publication. It is an opinion of the authors that because of applications, it makes sense to develop numerical methods, assuming that the question of uniqueness of the above inverse problem is addressed positively.

3 Nonlinear Integral Differential Equation Without the Unknown Coefficient

It follows from (6), (7) and the maximum principle that $w(x, s) > 0, \forall s \geq \underline{s}$. Consider the function $v = \ln w$. Since $x_0 \notin \overline{\Omega}$, then (6) and (8) lead to

$$\Delta v + |\nabla v|^2 = s^2 c(x) \quad \text{in } \Omega, \tag{11}$$

$$v(x, s) = \ln \varphi(x, s), \ \forall (x, s) \in \partial\Omega \times [\underline{s}, \overline{s}], \tag{12}$$

where $\varphi(x, s)$ is the Laplace transform (6) of the function $g(x, t)$. We eliminate the coefficient $c(x)$ from equation (11) via the differentiation with respect to s, since $\partial_s c(x) = 0$. To "isolate" the unknown coefficient in (11), introduce a new function

$$H(x, s) = \frac{v}{s^2}. \tag{13}$$

It follows from (13) and (9) that

$$D_x^\alpha(H) = O\left(\frac{1}{s}\right), D_x^\alpha D_s(H) = O\left(\frac{1}{s^2}\right), s \to \infty. \tag{14}$$

By (11)

$$\Delta H + s^2 (\nabla H)^2 = c(x). \tag{15}$$

Denote

$$q(x,s) = \partial_s H(x,s). \tag{16}$$

By (14) and (16)

$$H(x,s) = -\int_s^\infty q(x,\tau)\, d\tau.$$

We truncate this integral as

$$H(x,s) \approx -\int_s^{\overline{s}} q(x,\tau)\, d\tau + V(x,\overline{s}), \tag{17}$$

where $\overline{s} > s_0$ is a large number and

$$V(x,\overline{s}) \approx H(x,\overline{s}) = \frac{\ln w(x,\overline{s})}{\overline{s}^2}. \tag{18}$$

The number $\overline{s}$ should be chosen in numerical experiments. We call $V(x,\overline{s})$ the "tail", this function is unknown, and this is why we use " $\approx$ " here. By (14) the tail is small for large values of $\overline{s}$. In principle, therefore, one can set $V(x,\overline{s}) := 0$. However, our numerical experience shows that it would be better to update somehow the tail function in an iterative procedure. We call the updating procedure "iterations with respect to tails".

Thus, we obtain from (15)- (17) the following (approximate) integral nonlinear differential equation

$$\Delta q - 2s^2 \nabla q \cdot \int_s^{\overline{s}} \nabla q(x,\tau)\, d\tau + 2s \left[\int_s^{\overline{s}} \nabla q(x,\tau)\, d\tau\right]^2 \tag{19}$$

$$+2s^2 \nabla q \nabla V - 2s \nabla V \cdot \int_s^{\overline{s}} \nabla q(x,\tau)\, d\tau + 2s (\nabla V)^2 = 0$$

In addition, (12), (13) and (16) imply that the following Dirichlet boundary condition is given for the function q

$$q(x,s) = \psi(x,s),\ \forall (x,s) \in \partial\Omega \times [\underline{s},\overline{s}], \tag{20}$$

where

$$\psi(x,s) = \frac{\varphi_s}{\varphi s^2} - \frac{2\ln\varphi}{s^3}.$$

Suppose for a moment that the function q is approximated together with its derivatives $D_x^\alpha q, |\alpha| \le 2$. Then the corresponding approximation for the target coefficient can be found via (15) as

$$c(x) = \Delta H + \underline{s}^2 (\nabla H)^2, \tag{21}$$

where the function H is approximated via (17). Although any value of the pseudo frequency $s \in [\underline{s},\overline{s}]$ can be used in (21), but we found in our numerical experiments that the best value is $s := \underline{s}$. If integrals would be absent and the tail function would be known, then this would be the classic Dirichlet boundary value problem for the Laplace equation. However, the presence of integrals implies the nonlinearity and represents the main difficulty here.

Another obvious difficulty is that equation (19) has two unknown functions q and V. The reason why we can handle this difficulty is that we treat functions q and V differently: while we iteratively find approximations for q being sort of "restricted" only to equation (19), we find updates for V using solutions of forward problems (1), (2) and the formula (18). In those forward problems corresponding approximations for the unknown coefficient c, obtained from (21), are used.

4 A Sequence of Elliptic Dirichlet Boundary Value Problems

We approximate the function $q(x,s)$ as a piecewise constant function with respect to the pseudo frequency s. That is, we assume that there exists a partition $\underline{s} = s_N < s_{N-1} < ... < s_1 < s_0 = \overline{s}, s_{i-1} - s_i = h$ of the interval $[\underline{s}, \overline{s}]$ with a sufficiently small grid step size h such that $q(x,s) = q_n(x)$ for $s \in (s_n, s_{n-1})$. Hence,

$$\int_s^{\overline{s}} \nabla q(x,\tau) d\tau = (s_{n-1} - s)\nabla q_n(x) + h\sum_{j=1}^{n-1} \nabla q_j(x), s \in (s_n, s_{n-1}). \tag{22}$$

We approximate the boundary condition (20) as a piecewise constant function,

$$q_n(x) = \overline{\psi}_n(x), x \in \partial\Omega, \tag{23}$$

where

$$\overline{\psi}_n(x) = \frac{1}{h}\int_{s_n}^{s_{n-1}} \psi(x,s)\, ds. \tag{24}$$

On each subinterval $(s_n, s_{n-1}], n \geq 1$ we assume that functions $q_j(x)$, $j = 1, ..., n-1$ are known. We obtain an approximate equation for the function $q_n(x)$. Then we multiply this equation by the Carleman Weight Function (CWF) of the form

$$C_{n,\lambda}(s) = e^{\lambda(s - s_{n-1})}, \quad s \in (s_n, s_{n-1}], \tag{25}$$

and integrate with respect to s over (s_n, s_{n-1}). In (24) $\lambda >> 1$ is a parameter, which should be chosen in numerical experiments. Theorem 6.1 provides a recipe for this choice. We obtain (see details in [5])

$$\begin{aligned} L_n(q_n) &:= \Delta q_n - A_{1,n}\left(h\sum_{i=1}^{n-1}\nabla q_i\right)\nabla q_n + A_{1n}\nabla q_n \nabla V - \varepsilon q_n \\ &= 2\frac{I_{1,n}}{I_0}(\nabla q_n)^2 - A_{2,n}h^2\left(\sum_{i=1}^{n-1}\nabla q_i(x)\right)^2 \\ &+ 2A_{2,n}\nabla V\left(h\sum_{i=1}^{n-1}\nabla q_i\right) - A_{2,n}(\nabla V)^2, n = 1, ..., N, \end{aligned} \tag{26}$$

where $I_0 := I_0(\lambda, h), A_{1,n} := A_{1,n}(\lambda, h), A_{2,n} := A_{2,n}(\lambda, h)$ are certain integrals involving the CWF. Thus, we have obtained the Dirichlet boundary value problem (23), (26) for a nonlinear elliptic PDE with the unknown function $q_n(x)$. In this system the tail function V is also unknown. An important observation is that

$$\frac{|I_{1,n}(\lambda, h)|}{I_0(\lambda, h)} \leq \frac{4\overline{s}^2}{\lambda}. \tag{27}$$

Therefore, by taking $\lambda >> 1$, we mitigate the influence of the nonlinear term with $(\nabla q_n)^2$ in (26), which enables us to solve a linear problem on each iterative step. We have added the term $-\varepsilon q_n$ to the left hand side of equation (26), where $\varepsilon > 0$ is a small parameter. We are doing this because, by the maximum principle, if a function $p(x)$ is the classical solution of the Dirichlet boundary value problem

$$L_n(p) - \varepsilon p = f(x) \text{ in } \Omega, p\mid_{\partial\Omega} = p_b(x),$$

then [22] (Chapter 3, §1)

$$\max_{\overline{\Omega}} |p| \leq \max\left[\max_{\partial\Omega} |p_b|, \varepsilon^{-1} \max_{\overline{\Omega}} |f|\right].$$

On the other hand, if $\varepsilon = 0$, then the analogous estimate would be worse because of the involvement of some constants depending on $\max_{\overline{\Omega}} |\nabla q_j|$. Therefore, it is anticipated that the introduction of the term $-\varepsilon q_n$ should provide a better stability of our process, and we have indeed observed this in our computations.

5 The Algorithm

The above considerations lead to the algorithm described in this section. In particular, we describe here our procedure for iterative updates of the tails. We refer to subsection 5.4 of [18] for the first procedure of this sort, which was applied to a linearized CIP. Below $C^{k+\alpha}\left(\overline{\Omega}\right)$ are Hőlder spaces, where $k \geq 0$ is an integer and $\alpha \in (0,1)$ [22]. Denote $|f|_{k+\alpha} = \|f\|_{C^{k+\alpha}\left(\overline{\Omega}\right)}, \forall\ f \in C^{k+\alpha}\left(\overline{\Omega}\right)$. Our algorithm reconstructs iterative approximations $c_{n,i}(x) \in C^{\alpha}\left(\overline{\Omega}\right)$ of the function $c(x)$ only inside the domain Ω. On the other hand, to iterate with respect to tails, we need to solve the forward problem (1), (2). To do this, we need to extend each function $c_{n,k}(x)$ outside of the domain Ω in such a way that the resulting function $\widehat{c}_{n,k} \in C^{\alpha}\left(\mathbb{R}^3\right)$, $\widehat{c}_{n,k} \geq d_1$ in Ω and $\widehat{c}_{n,k} = 2d_1$ outside of Ω. The corresponding procedure is rather standard and is described in section 5 of [5]. In this section we mention convergencies of certain "sub-procedures". Numerical specifications of corresponding convergence criteria are given in subsection 9.1.

In accordance with (17), (21) and (22) denote

$$H_{n,i}(x) = hq_{n,i} + h\sum_{j=1}^{n-1} q_j(x) + V_{n,i}(x), x \in \Omega, \tag{28}$$

$$c_{n,i}(x) = \Delta H_{n,i} + s_n^2 \left(\nabla H_{n,i}\right)^2, \tag{29}$$

where functions $q_j, q_{n,i}, V_{n,i}$ are defined in this section below. Here m_n is the number of iterations with respect to tails for the given n where $i = 1, ..., m_n$. In our algorithm we set

$$q_0 := 0, q_{1,1}^0 := 0, V_{1,1}(x) := V_{1,1}^0(x), \tag{30}$$

$$q_{n,1}^0 := q_{n-1}, V_{n,1}(x) := V_{n-1,m_{n-1}}(x), \text{for } n \geq 2, \tag{31}$$

where $V_{1,1}^0(x)$ is a certain starting value for the tail function, which is specified in subsection 9.1.

Step $n_1, n \geq 1$. Suppose that functions $q_1, ..., q_{n-1}, q_{n,1}^0 := q_{n-1} \in C^{2+\alpha}(\overline{\Omega}), c_{n-1} \in C^{\alpha}(\overline{\Omega})$ and the tail function $V_{n,1}(x, \overline{s}) \in C^{2+\alpha}(\overline{\Omega})$ are constructed, see (30), (31) . We

now construct the function $q_{n,1}$. To do this, we solve iteratively the following Dirichlet boundary value problems

$$\Delta q_{n,1}^k - A_{1n}\left(h\sum_{j=1}^{n-1}\nabla q_j\right)\cdot\nabla q_{n,1}^k - \varepsilon q_{n,1}^k + A_{1n}\nabla q_{n,1}^k\cdot\nabla V_{n,1} =$$

$$2\frac{I_{1n}}{I_0}\left(\nabla q_{n,1}^{k-1}\right)^2 - A_{2n}h^2\left(\sum_{j=1}^{n-1}\nabla q_j(x)\right)^2 \tag{32}$$

$$+2A_{2n}\nabla V_{n,1}\cdot\left(h\sum_{j=1}^{n-1}\nabla q_j(x)\right) - A_{2n}\left(\nabla V_{n,1}\right)^2, q_{n,1}^k \in C^{2+\alpha}\left(\overline{\Omega}\right), k=1,2,...,$$

$$q_{n,1}^k(x) = \overline{\psi}_n(x), x\in\partial\Omega. \tag{33}$$

We call these "iterations with respect to the nonlinear term". We iterate here until the process converges. Then we set

$$q_{n,1} = \lim_{k\to\infty} q_{n,1}^k \text{ in the } C^{2+\alpha}\left(\overline{\Omega}\right) \text{ norm.}$$

Next, we reconstruct an approximation $c_{n,1}(x), x\in\Omega$ for the unknown function $c(x)$ using the resulting function $q_{n,1}(x)$ and formulas (28), (29) at $i=1$. Hence, $c_{n,1}\in C^{\alpha}\left(\overline{\Omega}\right)$. Assume that $c_{n,1}(x)\geq d_1$ in Ω. Construct the function $\widehat{c}_{n,1}(x)\in C^{\alpha}\left(\mathbb{R}^3\right)$. Next, solve the forward problem (1), (2) with $c(x):=\widehat{c}_{n,1}(x)$. We obtain the function $u_{n,1}(x,t)$. Calculate the Laplace transform (6) of this function and obtain the function $w_{n,1}(x,\overline{s})$ this way. Next, following (18), we set for $x\in\Omega$

$$V_{n,2}(x,\overline{s}) = \frac{1}{\overline{s}^2}\ln w_{n,1}(x,\overline{s}) \in C^{2+\alpha}\left(\overline{\Omega}\right). \tag{34}$$

Step n_i, $i\geq 2, n\geq 1$. Suppose that functions $q_{n,i-1}, V_{n,i}(x,\overline{s})\in C^{2+\alpha}\left(\overline{\Omega}\right)$ are constructed. We now iterate with respect to the tail only. That is, we solve the boundary value problem

$$\Delta q_{n,i} - A_{1n}\left(h\sum_{j=1}^{n-1}\nabla q_j\right)\cdot\nabla q_{n,i} - \varepsilon q_{n,i} + A_{1n}\nabla q_{n,i}\cdot\nabla V_{n,i}$$

$$= 2\frac{I_{1n}}{I_0}\left(\nabla q_{n,i-1}\right)^2 - A_{2n}h^2\left(\sum_{j=1}^{n-1}\nabla q_j(x)\right)^2 \tag{35}$$

$$+ 2A_{2n}\nabla V_{n,i}\cdot\left(h\sum_{j=1}^{n-1}\nabla q_j(x)\right) - A_{2n}\left(\nabla V_{n,i}\right)^2,$$

$$q_{n,i}(x) = \overline{\psi}_n(x), x\in\partial\Omega. \tag{36}$$

Having the function $q_{n,i}$, we reconstruct the next approximation $c_{n,i}\in C^{\alpha}(\overline{\Omega})$ for the target coefficient using (28), (29), and, assuming that $c_{n,i}(x)\geq d_1$ in Ω, construct the function $\widehat{c}_{n,i}\in C^{\alpha}(\mathbb{R}^3)$. Next, we solve the forward problem (1), (2) with $c(x):=\widehat{c}_{n,i}(x)$, calculate the Laplace transform (6) and update the tail as in (34), where $(w_{n,1}, V_{n,2})$ is replaced with

$(w_{n,i}, V_{n,i+1})$. We iterate with respect to i until convergence occurs at the step $i := m_n$. Then we set

$$q_n := q_{n,m_n} \in C^{2+\alpha}\left(\overline{\Omega}\right), c_n := c_{n,m_n} \in C^{\alpha}\left(\overline{\Omega}\right),$$

$$V_{n+1,1}(x,\overline{s}) = \frac{1}{\overline{s}^2} \ln w_{n,m_n}(x,\overline{s}) \in C^{2+\alpha}\left(\overline{\Omega}\right). \tag{37}$$

If functions $c_n(x)$ did not yet converge, then we proceed with Step $(n+1)_1$, provided that $n < \overline{N}$, where $\overline{N}$ is a prescribed iteration number, $\overline{N} \leq N$, see Theorem 6.1. However, if either functions $c_n(x)$ converged, or $n = \overline{N}$, then we stop. It follows from (34) that in principle, to update the tail, one can solve the problem (7), (8) for $s = \overline{s}$ instead of the problem (1), (2). However, our computational experience shows that it is better to proceed via solving the problem (1), (2) and calculating the Laplace transform then. We do not yet have an explanation for this.

6 Global Convergence

By the concept of Tikhonov for ill-posed problems [30], which we follow, one should assume first that there exists an "ideal" exact solution of an ill-posed problem with the exact data. Next, one should assume the presence of an error of the level ζ in the data, where $\zeta > 0$ is a small parameter. Suppose that an approximate solution is constructed for each sufficiently small ζ. This solution is called a "regularized solution", if it tends to the exact solution as $\zeta \to 0$.

6.1 Exact solution

First, we introduce the definition of the exact solution. We assume that there exists a coefficient $c^*(x) \in [2d_1, 2d_2]$ satisfying condition (4), and this function is an exact solution of our Inverse Problem with the exact data in $g^*(x,t)$ in (8). The Laplace transform (6) of the function $g^*(x,t)$ leads to the exact function $\varphi^*(x,s) = w^*(x,s), \forall (x,s) \in \partial\Omega \times [\underline{s}, \overline{s}]$. Here the function $w^*(x,s) \in C^{2+\alpha}\left(\mathbb{R}^3 \diagdown \{|x - x_0| < \gamma\}\right), \forall \gamma > 0, \forall s \geq \underline{s}$ is the solution of the forward problem (7), (8) with $c(x) := c^*(x)$. Also, let

$$H^*(x,s) = \frac{\ln[w^*(x,s)]}{s^2},\ q^*(x,s) = \frac{\partial H^*(x,s)}{\partial s},\ V^*(x,\overline{s}) = H^*(x,\overline{s}).$$

The function q^* satisfies an obvious analogue of equation (19) with the boundary condition (see (20)

$$q^*(x,s) = \psi^*(x,s), (x,s) \in \partial\Omega \times [\underline{s}, \overline{s}], \tag{38}$$

where

$$\psi^*(x,s) = \frac{1}{\varphi^* s^2} \cdot \frac{\partial \varphi^*}{\partial s} - \frac{2 \ln \varphi^*}{s^3}.$$

Definition. We call the function $q^*(x,s)$ *the exact solution* of the problem (19), (20) with the *exact boundary* condition $\psi^*(x,s)$.

Hence,

$$q^*(x,s) \in C^{2+\alpha}\left(\overline{\Omega}\right) \times C^{\infty}[\underline{s}, \overline{s}]. \tag{39}$$

We now follow (22)-(26). First, we approximate functions $q^*(x,s)$ and $\psi^*(x,s)$ via piecewise constant functions with respect to $s \in [\underline{s}, \overline{s}]$. Let

$$q_n^*(x) = \frac{1}{h} \int\limits_{s_n}^{s_{n-1}} q^*(x,s)\, ds,\ \overline{\psi}_n^*(x) = \frac{1}{h} \int\limits_{s_n}^{s_{n-1}} \psi^*(x,s)\, ds. \tag{40}$$

Then

$$q^*(x,s) = q_n^*(x) + Q_n(x,s), \psi^*(x,s) = \overline{\psi}_n^*(x) + \Psi_n(x,s), s \in [s_n, s_{n-1}],$$

where by (6.1) functions Q_n, Ψ_n are such that

$$|Q_n(x,s)|_{2+\alpha} \leq C^* h, |\Psi_n(x,s)|_{2+\alpha} \leq C^* h, n = 1, ..., N, \text{ for } s \in [s_n, s_{n-1}], \tag{41}$$

where the constant $C^* = C^*\left(\|q^*\|_{C^{2+\alpha}(\overline{\Omega}) \times C^1[\underline{s}, \overline{s}]}\right) > 0$ depends only on the $C^{2+\alpha}(\overline{\Omega}) \times C^1[\underline{s}, \overline{s}]$ norm of the function $q^*(x,s)$. Hence, we can assume that

$$\max_{1 \leq n \leq N} |q_n^*|_{2+\alpha} \leq C^*. \tag{42}$$

Without a loss of generality, we assume that

$$C^* \geq 1. \tag{43}$$

By the Tikhonov concept, the constant C^* should be known a priori. By Lemma 2.1, it is reasonable to assume that C^* is independent on $\overline{s}$, although we do not use this assumption. By (40)

$$q_n^*(x) = \overline{\psi}_n^*(x), x \in \partial\Omega. \tag{44}$$

Hence we obtain the following analogue of the equation (26) from (39)

$$\begin{aligned} &\Delta q_n^* - A_{1,n}\left(h\sum_{i=1}^{n-1} \nabla q_i^*(x)\right)\nabla q_n^* + A_{1,n}\nabla q_n^* \nabla V^* \\ &= 2\frac{I_{1,n}}{I_0}(\nabla q_n^*)^2 - A_{2,n}h^2\left(\sum_{i=1}^{n-1}\nabla q_i^*(x)\right)^2 \\ &+2A_{2,n}\nabla V^*\left(h\sum_{i=1}^{n-1}\nabla q_i^*(x)\right) - A_{2,n}|\nabla V^*|^2 + F_n(x,h,\lambda), \end{aligned} \tag{45}$$

where the function $F_n(x,h,\lambda) \in C^\alpha(\overline{\Omega})$ and

$$\max_{\lambda h \geq 1} |F_n(x,h,\lambda)|_\alpha \leq C^* h. \tag{46}$$

We also assume that the function $g(x,t)$ in (8) is given with an error. This naturally produces an error in the function $\psi(x,s)$ in (20). An additional error is introduced due to the averaging in (24). Hence, it is reasonable to assume that

$$\left\|\overline{\psi}_n^*(x) - \overline{\psi}_n(x)\right\|_{C^{2+\alpha}(\partial\Omega)} \leq C^*(\sigma + h), \tag{47}$$

where $\sigma > 0$ is a small parameter characterizing the level of the error in the data $\psi(x,s)$. The parameter h can also be considered as a part of the error in the data, since we have replaced a smooth s-dependent function with a piecewise constant one.

6.2 Convergence theorem

First, we reformulate the Schauder theorem in a simplified form, which is convenient for our case, see Chapter 3, §1 in [22] for this theorem. Assuming that

$$\overline{s} > 1,\ \lambda h \geq 1, \tag{48}$$

it was shown in [5] that

$$\max_{1 \leq n \leq N} \{|A_{1,n}| + |A_{2,n}|\} \leq 8\overline{s}^2. \tag{49}$$

Introduce the positive constant $M^* = M^*\left(\|q^*\|_{C^{2+\alpha}(\overline{\Omega}) \times C^1[\underline{s},\overline{s}]}, \overline{s}\right) = M^*\left(C^*, \overline{s}\right)$ by

$$M^* = 2C^* \max\left(8\overline{s}^2, \max_{1 \leq n \leq N} \{|A_{1,n}| + |A_{2,n}|\}\right). \tag{50}$$

Hence, (49) and (50) imply that

$$M^* = 16C^*\overline{s}^2. \tag{51}$$

Consider the Dirichlet boundary value problem

$$\Delta u + \sum_{j=1}^{3} b_j(x) u_{x_j} - d(x)u = f(x),\ x \in \Omega,$$

$$u\,|_{\partial\Omega} = g(x) \in C^{2+\alpha}(\partial\Omega).$$

Assume that the following conditions are satisfied

$$b_j, d, f \in C^{\alpha}\left(\overline{\Omega}\right), d(x) \geq 0;\ \max\left(|b_j|_{\alpha}, |d|_{\alpha}\right) \leq 1.$$

By the Schauder theorem, there exists unique solution $u \in C^{2+\alpha}\left(\overline{\Omega}\right)$ of this boundary value problem, and with a constant $K = K(\Omega) > 0$ depending only on the number the domain Ω the following estimate holds

$$|u|_{2+\alpha} \leq K\left[\|g\|_{C^{2+\alpha}(\partial\Omega)} + |f|_{\alpha}\right]. \tag{52}$$

In the formulation of Theorem 6.1 we provide estimates (56)-(61) via M^* and also use (51) to obtain estimates via $\overline{s}$. This formulation is almost the same as one in [5]. Note that the definition of the norm in the space $C^{\alpha}(\overline{\Omega})$ implies that

$$|f_1 f_2|_{\alpha} \leq |f_1|_{\alpha} |f_2|_{\alpha},\ \forall f_1, f_2 \in C^{\alpha}\left(\overline{\Omega}\right). \tag{53}$$

Theorem 6.1. *Let $\Omega \subset \mathbb{R}^3$ be a convex bounded domain with the boundary $\partial\Omega \in C^3$. Suppose that (42)-(44) and (46)-(48) hold. Let the exact coefficient $c^*(x)$ satisfies (4) and $c^* \in [2d_1, 2d_2], c^*(x) = 2d_1$ for $x \in \mathbb{R}^3 \setminus \Omega$, where numbers $d_1, d_2 > 0$ are given. For any function $c(x) \in C^{\alpha}\left(\mathbb{R}^3\right)$ such that $c(x) \geq d_1$ in Ω and $c(x) = 2d_1$ in $\mathbb{R}^3 \setminus \Omega$ consider the solution $u(x,t)$ of the Cauchy problem (1), (2). Let $w_c(x,s) \in C^{2+\alpha}\left(\mathbb{R}^3 \setminus \{|x - x_0| < \gamma\}\right), \forall \gamma > 0$ be the Laplace transform (6) of $u(x,t)$ and $V_c(x) = \overline{s}^{-2} \ln w_c(x,\overline{s}) \in C^{2+\alpha}\left(\overline{\Omega}\right)$ be the corresponding tail function. Suppose that the cut-off pseudo frequency $\overline{s}$ is so large that both for $c^*(x)$ and any such function $c(x)$ the following estimates hold (see (14))*

$$|V^*|_{2+\alpha} \leq \xi, |V_c|_{2+\alpha} \leq \xi, \tag{54}$$

where $\xi \in (0,1)$ is a sufficiently small number. Let $V_{1,1}(x,\overline{s}) \in C^{2+\alpha}\left(\overline{\Omega}\right)$ be the initial tail function and let (see (30))

$$|V_{1,1}|_{2+\alpha} \leq \xi. \tag{55}$$

Denote $\eta := 2(h+\sigma+\xi+\varepsilon)$. Let $K = K(\Omega) > 0$ be the constant of the Schauder theorem in (52) and $\overline{N} \leq N$ be the total number of functions q_n calculated by the above algorithm. Suppose that the number $\overline{N} = \overline{N}(h)$ is connected with the step size h via $\overline{N}(h)\, h = \beta$, where the constant $\beta > 0$ is independent on h. Let β be so small that

$$\beta \leq \min\left(\frac{2}{7}, \frac{1}{16^2 K C^* \overline{s}^4}, \frac{1}{16^2 C^* \overline{s}^4}\right) \leq \min\left(\frac{2}{7}, \frac{1}{16 K M^* \overline{s}^2}, \frac{C^*}{(M^*)^2}\right). \tag{56}$$

In addition, let the number η and the parameter λ of the CWF satisfy the following estimates

$$\eta \leq \eta_0(\Omega, M^*, d_1) = \eta_0\left(\Omega, \|q^*\|_{C^{2+\alpha}\left(\overline{\Omega}\right)\times C^1[\underline{s},\overline{s}]}, d_1, \overline{s}\right)$$

$$= \min\left(\frac{1}{2}, \frac{1}{4K}, \frac{d_1}{32\cdot 16 C^* \overline{s}^4}, \frac{1}{8 C^* \overline{s}^2}\right) = \min\left(\frac{1}{2}, \frac{1}{4K}, \frac{d_1}{32 M^* \overline{s}^2}, \frac{2}{M^*}\right), \tag{57}$$

$$\lambda \geq \lambda_0(C^*, K, \overline{s}, \eta) = \max\left(16^4 (C^*)^4 \overline{s}^8, 6\cdot 16^2 (C^*)^2 K \overline{s}^4, \frac{1}{\eta^2}\right). \tag{58}$$

Then for every integer $n \in \left[1, \overline{N}\right]$ the following estimates hold

$$\left|q_{n,1}^k - q_n^*\right|_{2+\alpha}, |q_{n,i} - q_n^*|_{2+\alpha} \leq 2KM^*\left(\frac{1}{\sqrt{\lambda}} + 3\eta\right) = 32 C^* K \overline{s}^2\left(\frac{1}{\sqrt{\lambda}} + 3\eta\right), \tag{59}$$

$$\left|q_{n,1}^k\right|_{2+\alpha}, |q_{n,i}|_{2+\alpha} \leq 2C^*, \tag{60}$$

$$|c_{n,i} - c^*|_\alpha \leq 8 M^* \overline{s}^2\left(\frac{1}{\sqrt{\lambda}} + 3\eta\right) = 128 C^* \overline{s}^4\left(\frac{1}{\sqrt{\lambda}} + 3\eta\right). \tag{61}$$

In addition, functions $c_{n,i}(x) \geq d_1$ in Ω and $\widehat{c}_{n,k}(x) \geq d_1$ in $\mathbb{R}^3$.

Proof. The major part of the proof is the same as one in [5]. The only thing we prove now is the convergence of functions $q_{n,1}^k$ for $k \to \infty$, because it was not proven in [5].The idea of the proof is to consider the differences $\widetilde{q}_{n,i}^k = q_{n,i}^k - q_n^*, \widetilde{q}_{n,i} = q_{n,i} - q_n^*$, obtain Dirichlet boundary value problems for linear elliptic equations for them via subtracting (45) from either (32) or (35), and (44) from either (33) or (36), and then sequentially use the estimate (52) of the Schauder theorem to estimate norms $\left|\widetilde{q}_{n,i}^k\right|_{2+\alpha}, |\widetilde{q}_{n,i}|_{2+\alpha}$ from the above. In doing so, one needs to estimate differences of tails $\widetilde{V}_{n,k} = V_{n,k} - V^*$ using (37), (54) and (55), as $\left|\widetilde{V}_{n,k}\right|_{2+\alpha} \leq 2\xi \leq \eta$.

It follows from [5] that given the number n, estimates (59) and (60) for $\left|q_{n,1}^k - q_n^*\right|_{2+\alpha}, \left|q_{n,1}^k\right|_{2+\alpha}$ can be proven, using the above outlined idea, without the proof of convergence of functions $q_{n,1}^k, k \to \infty$. Hence, we assume now that these estimates are valid. Consider for example the case $n = 1$, since other cases are similar. Let $m, r > 2$ be two positive integers. Denote $a_{m,r} = q_{1,1}^m - q_{1,1}^r$. Setting in (32), (33) $n = 1, k := m$, then $k := r$ and subtracting two resulting equations, we obtain

$$\Delta a_{m,r} - \varepsilon a_{m,r} + A_{1,1}\nabla a_{mr}\cdot\nabla V_{1,1} = 2\frac{I_{1,1}}{I_0}\nabla a_{m-1,r-1}\cdot\left(\nabla q_{1,1}^{m-1} + \nabla q_{1,1}^{r-1}\right), \tag{62}$$

$$a_{m,r}\,|_{\partial\Omega} = 0. \tag{63}$$

For the vector function $f = (f_1, f_2, f_3), f_i \in C^\alpha\left(\overline{\Omega}\right)$ denote

$$|f|_\alpha = \left[\sum_{j=1}^{3} \left|f_{x_j}\right|_\alpha^2\right]^{1/2}.$$

Hence, $|\nabla a_{m-1,r-1}|_\alpha \leq \sqrt{3}\,|a_{m-1,r-1}|_{2+\alpha}$ and by (60) $\left|\nabla q_{1,1}^{m-1} + \nabla q_{1,1}^{r-1}\right|_\alpha \leq 4\sqrt{3}C^*$. Hence, (27) and (53) lead to

$$\left|2\frac{I_{1,1}}{I_0}\nabla a_{m-1,r-1} \cdot \left(\nabla q_{1,1}^{m-1} + \nabla q_{1,1}^{r-1}\right)\right|_\alpha \leq \frac{100C^*\overline{s}^2}{\lambda}\,|a_{m-1,r-1}|_{2+\alpha}\,. \tag{64}$$

Applying (52) to (62), (63), and taking into account (58) and (64), we obtain

$$|a_{m,r}|_{2+\alpha} \leq \frac{100KC^*\overline{s}^2}{\lambda}\,|a_{m-1,r-1}|_{2+\alpha} \leq \frac{100}{6\cdot 16^2 C^*\overline{s}^2}\,|a_{m-1,r-1}|_{2+\alpha}\,. \tag{65}$$

Since by (43) and (48) $C^*\overline{s}^2 > 1$, then (65) implies that

$$|a_{m,r}|_{2+\alpha} \leq \frac{1}{10}\,|a_{m-1,r-1}|_{2+\alpha}\,. \tag{66}$$

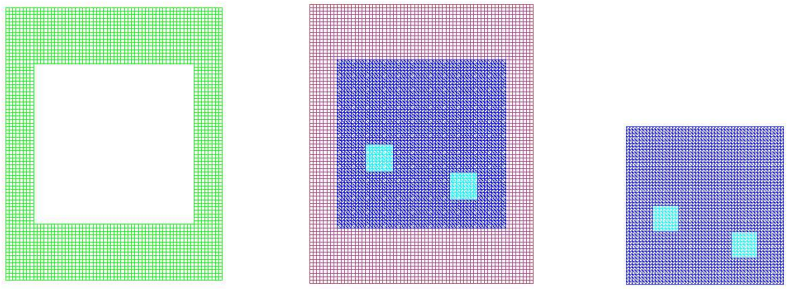
It follows from (66) that the sequence $\left\{q_{1,1}^k\right\}_{k=1}^\infty$ satisfies the Cauchy convergence criterion. $\square$.

Remarks:

1. It often happens in the computational practice of ill-posed problems that theoretical estimates of convergence theorems are more pessimistic than ones obtained in numerical studies, and also some discrepancies between analytical results and their numerical implementations often occur. Our computational experience tells us that this is exactly our case in reference to estimates (56)- (61). It seems to be at the first glance that because of (61), one can stop the iterative process at $n = 1$. However, our numerical experience shows that this way one cannot obtain good images. Hence, we use in our computations a stopping rule, which is different from (61). Actually, we do not use the $C^{2+\alpha}(\overline{\Omega})$ norm to verify convergence, because it is rather complicated in the computational practice to consider this norm and also because all norms in finite dimensional spaces are equivalent, and we work in a finite dimensional space of finite elements in our computations. In addition, we have used the $\delta-$ function in (1) and the whole space $\mathbb{R}^3$ only for the sake of a convenient formulation of the asymptotic behavior (9). In our computations we use the plane wave and a bounded domain G for the solution of the forward problem. Other main discrepancies between our theory and the computational implementation are listed in subsection 7.2 of [5]. In particular, it is stated there that we verify the asymptotic behavior at $s \to \infty$ computationally.

2. Truncating integrals at a high pseudo frequency $\overline{s}$ is a natural thing to do, because one routinely truncates high frequencies in physics and engineering. By truncating integrals, we actually come up with a different, although a quite reasonable mathematical model.

3. One of the back bones of the theory of ill-posed problems is that the number of iterations can be chosen as a regularization parameter, see, e.g., page 157 of [16]. Therefore, we have a vector $(\overline{s}, \overline{N}, m_1, ..., m_{\overline{N}})$ of regularization parameters, see details about their choice in subsection 8.2. Setting $\overline{N}(h)\,h = \beta = const. > 0$ is in an agreement with, e.g., Lemma 6.2 on page 156 of [16], since this lemma shows a connection between the error in the data and the number of iterations (that lemma is proven for a different algorithm). The



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

Figure 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 $x \in G \diagdown \Omega$.

number β is small because our algorithm is originated by equation (19), which contains nonlinear terms with s-integrals of the Volterra type. It well known that in general solutions of nonlinear integral equations of the Volterra type can be estimated only on sufficiently small intervals.

7 Computations of the Forward Problem

In this paper we work with the computationally simulated data. That is, the data are generated by computing the forward problem (68) with the given function $c(x)$. To solve the forward problem, we use the hybrid FEM/FDM method described in [11]. The computational domain in all our tests $G = G_{FEM} \cup G_{FDM}$ is set as $G = [-4.0, 4.0] \times [-5.0, 5.0]$. This domain is split into a finite element domain $G_{FEM} := \Omega = [-3.0, 3.0] \times [-3.0, 3.0]$ and a surrounding domain G_{FDM} with a structured mesh, see Figure 1. The space mesh in Ω consists of triangles and in G_{FDM} - of squares with the mesh size $\tilde{h} = 0.125$ in the overlapping regions. At the top and bottom boundaries of G we use first-order absorbing boundary conditions [17] which are exact in this particular case since the plane wave is initialized in normal direction into G in all our tests. At the lateral boundaries, mirror boundary conditions allow us to assume an infinite space domain in the lateral direction.

The forward problem is computed in the domain $G \subset \mathbb{R}^2$ (Figure 1). The coefficient $c(x)$ is unknown only in domain $\Omega \subset G$ and

$$c(x) = 1 \text{ in } G \diagdown \Omega. \tag{67}$$

The trace of the solution of the forward problem is recorded at the boundary $\partial\Omega$. Next, the coefficient $c(x)$ is "forgotten", and our goal is to reconstruct this coefficient for $x \in \Omega$ from the data $\varphi(x, s)$. The boundary of the domain G is $\partial G = \partial G_1 \cup \partial G_2 \cup \partial G_3$. Here, ∂G_1 and ∂G_2 are respectively top and bottom sides of the largest domain of Figure 1 and ∂G_3 is the union of left and right sides of this domain. In our first test the forward problem

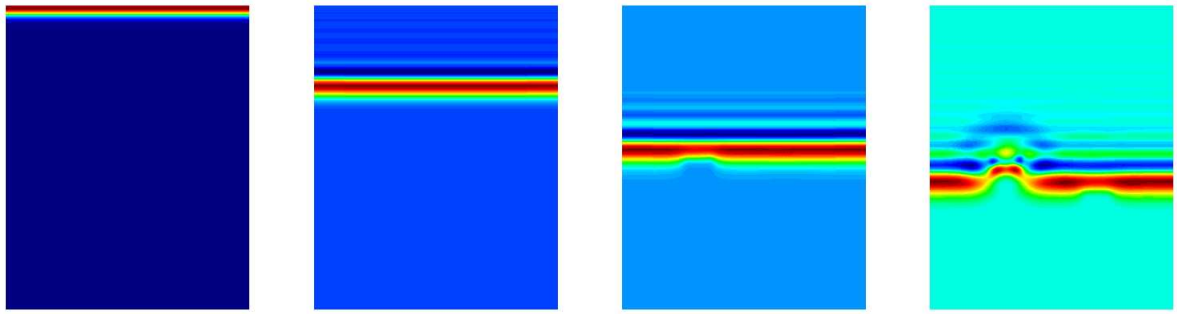


Figure 2: Test 1: Isosurfaces of the simulated exact solution to the forward problem (68) at different times with a plane wave initialized at the top boundary.

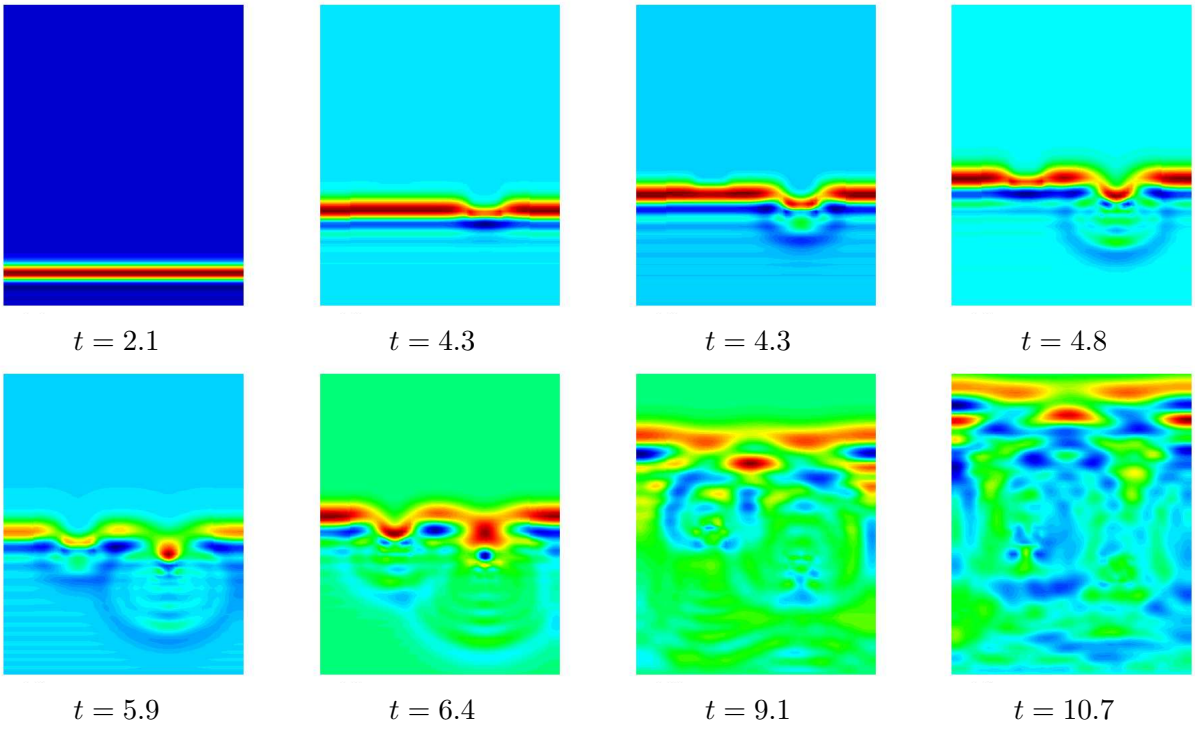


Figure 3: Test 2: Isosurfaces of the simulated exact solution to the forward problem (68) with a plane wave initialized at the bottom boundary.

clearpage
clearpage
is

$$
\begin{aligned}
c(x)\frac{\partial^2 u}{\partial t^2} - \triangle u = 0, \quad & \text{in } G \times (0,T), \\
u(\cdot,0) = 0, \ \frac{\partial u}{\partial t}(\cdot,0) = 0, \ & \text{in } G, \\
\partial_n u\big|_{\partial G_1} = f(t), \ & \text{on } \partial G_1 \times (0,t_1], \\
\partial_n u\big|_{\partial G_1} = \partial_t u, \ & \text{on } \partial G_1 \times (t_1,T), \\
\partial_n u\big|_{\partial G_2} = \partial_t u, \ & \text{on } \partial G_2 \times (0,T), \\
\partial_n u\big|_{\partial G_3} = 0, \ & \text{on } \partial G_3 \times (0,T),
\end{aligned}
\tag{68}
$$

where T is the final time and $f(t)$ is the plane wave defined as

$$
f(t) = \frac{(\sin(\overline{s}t - \pi/2) + 1)}{10}, \ 0 \le t \le t_1 := \frac{2\pi}{\overline{s}}, T = 17.8 t_1.
$$

Thus, the plane wave is initialized at the top boundary ∂G_1 and propagates into G for $t \in (0,t_1]$. First order absorbing boundary conditions [17] are used on $\partial G_1 \times (t_1,T]$ and $\partial G_2 \times (0,T]$, and the Neumann boundary condition is used on the bottom boundary ∂G_3. In the second test we consider the case when the plane wave is initialized at the bottom boundary and use the Neumann boundary condition at the top boundary. In the integral (6) of the Laplace transform we integrate for $t \in (0,T)$.

8 The Adaptivity Technique

In this section we describe the adaptivity technique, for the sake of completeness, as well as for the sake of the above mentioned rigorous explanation of the meaning of the ψ function. However, we are not proving the estimate (95) here for the Frechet derivative of the Lagrangian and instead refer for details to [7, 8, 9].

To use the adaptivity technique, we formulate the inverse problem for the boundary value problem (68) as an optimization problem, where we seek the unknown coefficient $c(x)$, which gives the solution of the boundary value problem (68) for the function $u(x,t)$ with the best least squares fit to the time domain observations $g(x,t)$, see (5). Denote $Q_T = \Omega \times (0,T), S_T = \partial\Omega \times (0,T)$. Our goal now is to find the function $c(x)$ which minimizes the Tikhonov functional

$$E(u,c) = \frac{1}{2}\int\limits_{S_T} (u\mid_{S_T} - g(x,t))^2 d\sigma dt + \frac{1}{2}\gamma \int\limits_{\Omega} (c-c_0)^2 \ \ dx, \tag{69}$$

where γ is the regularization parameter and c_0 is an initial guess for the unknown coefficient c. On the first step of the adaptivity we take the same mesh as one we have used for the globally convergent method. The first guess $c_0 = c_0(x)$ is also taken the one, which was obtained on the globally convergent stage. On each follow up step of the adaptivity when mesh refinement is used, the function $c_0(x)$ is taken from the previous step. In doing so, values of that function are linearly interpolated from the coarser grid on the finer grid. Note that since $c(x) = 1$ in the domain $G \diagdown \Omega$, then given the function $g(x,t) = u\mid_{\partial\Omega}$, one can uniquely determine the function $u(x,t)$ for $(x,t) \in (G \diagdown \Omega) \times (0,T)$ as the solution of the boundary value problem for equation (68) for with boundary conditions on both boundaries ∂G and $\partial\Omega$. Hence, one can uniquely determine the function $p(x,t)$,

$$\frac{\partial u}{\partial n}\mid_{S_T} = p(x,t). \tag{70}$$

Since we deal with computationally simulated data, in our computations both functions $p(x,t)$ and $g(x,t)$ are calculated from the solution of the forward problem (68) with the correct value of the coefficient $c(x)$.

Denote

$$\begin{aligned}
H_u^2(Q_T) &= \{f \in H^2(Q_T) : f(x,0) = f_t(x,0) = 0\},\\
H_u^1(Q_T) &= \{f \in H^1(Q_T) : f(x,0) = 0\},\\
H_\varphi^2(Q_T) &= \{f \in H^2(Q_T) : f(x,T) = f_t(x,T) = 0\},\\
H_\varphi^1(Q_T) &= \{f \in H^1(Q_T) : f(x,T) = 0\},\\
U &= H_u^2(Q_T) \times H_\varphi^2(Q_T) \times C^2(\overline{\Omega}),\\
\bar{U} &= H_u^1(Q_T) \times H_\varphi^1(Q_T) \times L_2(\Omega),\\
\bar{U}^1 &= L_2(Q_T) \times L_2(Q_T) \times L_2(\Omega)
\end{aligned} \tag{71}$$

where all functions are real valued. Hence, $U \subset \bar{U} \subset \bar{U}^1$ as sets, U is dense in $\bar{U}$ and $\bar{U}$ is dense in $\bar{U}^1$. Also denote $((\cdot,\cdot))$ the inner product in $\bar{U}^1$ and $[\cdot]$ the norm generated by this product.

To solve the problem of the minimization of the functional (69), we introduce the Lagrangian

$$L(v) = E(u,c) + \int\limits_{Q_T} \varphi \cdot (cu_{tt} - \Delta u)\, dxdt, \forall \varphi \in H_\varphi^2(Q_T), \tag{72}$$

where $\varphi \in H_{\varphi}^{2}(Q_T)$ is the Lagrange multiplier and $v = (u, \varphi, c) \in U$. Since the function u solves equation (68) then $L(v) = E(u, c)$. This is because the second term in $L(v)$ is zero. Integration by parts and (72) leads to

$$L(v) = E(u,c) - \int\limits_{Q_T} c(x) u_t \varphi_t dx dt + \int\limits_{Q_T} \nabla u \nabla \varphi dx dt - \int\limits_{S_T} p\varphi d\sigma dt. \tag{73}$$

We search for a stationary point of the functional $L(v), v \in U$ satisfying

$$L'(v)(\overline{v}) = 0, \quad \forall \overline{v} = (\bar{u}, \bar{\varphi}, \bar{c}) \in \bar{U} \tag{74}$$

where $L'(v)(\cdot)$ is the Frechet derivative of L at the point v. To find the Frechet derivative $L'(v)(\overline{v})$, consider $L(v+\overline{v}) - L(v)\ \forall \overline{v} \in \bar{U}$ and single out the linear, with respect to $\overline{v}$, part of this expression. Hence, we obtain from (73) and (74)

$$L'(v)(\overline{v}) = \int\limits_{\Omega} \bar{c} \left[\gamma (c - c_0) - \int\limits_0^T u_t \varphi_t dt \right] dx - \int\limits_{Q_T} c(x) \left(\varphi_t \overline{u}_t + u_t \overline{\varphi}_t \right) dx dt \tag{75}$$

$$+ \int\limits_{Q_T} \left(\nabla u \nabla \overline{\varphi} + \nabla \overline{u} \nabla \varphi \right) - \int\limits_{S_T} p \overline{\varphi} d\sigma dt = 0, \forall \overline{v} = (\overline{u}, \overline{\varphi}, \overline{c}) \in \bar{U}.$$

Integration by parts in (75) leads to

$$L'(v)(\overline{v}) = \int\limits_{\Omega} \bar{c} \left[\gamma (c - c_0) - \int\limits_0^T u_t \varphi_t dt \right] dx + \int\limits_{Q_T} \bar{\varphi} \left(c u_{tt} - \Delta u \right) dx dt \tag{76}$$

$$+ \int\limits_{Q_T} \bar{u} \left(c \varphi_{tt} - \Delta \varphi \right) dx dt + \int\limits_{S_T} \bar{u} \left[(u - g) - \partial_n \varphi \right] d\sigma dt, \forall \overline{v} = (\overline{u}, \overline{\varphi}, \overline{c}) \in \bar{U}.$$

Hence (74) and (75) imply that every integral term in formula (76) equals zero. We obtain that if $(u, \varphi, c) = v \in U$ is a minimizer of the Lagrangian $L(v)$ in (73), then

$$c u_{tt} - \triangle u = 0, (x,t) \in Q_T, \tag{77}$$

$$u(x,0) = u_t(x,0) = 0, \tag{78}$$

$$\partial_n u \mid_{S_T} = p(x,t); \tag{79}$$

$$c \varphi_{tt} - \triangle \varphi = 0, (x,t) \in Q_T, \tag{80}$$

$$\varphi(x,T) = \varphi_t(x,T) = 0, \tag{81}$$

$$\frac{\partial \varphi}{\partial n} \mid_{S_T} = (u - g)(x,t), (x,t) \in S_T; \tag{82}$$

$$\gamma (c - c_0) - \int_0^T u_t \varphi_t \, dt = 0, x \in \Omega. \tag{83}$$

The boundary value problem (80)-(82) should be solved backwards in time. Uniqueness and existence theorems for initial boundary value problems (77)-(79) and (80)-(82), including the case of weak $H_u^1(Q_T)$ and $H_{\varphi}^1(Q_T)$ solutions, can be found in Chapter 4 of [23]. We minimize $L(v)$ in an iterative process via solving on each step boundary value problems

(77)-(79) and (80)-(82). We find weak solutions of problems (77)-(79), (80)-(82) via the FEM.

To formulate the FEM for boundary value problems (77)-(79) and (80)-(82) we introduce finite element spaces $W_h^u \subset H_u^1(Q_T)$ and $W_h^\varphi \subset H_\varphi^1(Q_T)$ for functions u and φ respectively. These spaces consist of continuous piecewise linear functions in space and time satisfying initial conditions $u(x,0) = 0$ for $u \in W_h^u$ and $\varphi(x,T) = 0$ for $\varphi \in W_h^\varphi$. We also introduce the finite element space $V_h \subset L_2(\Omega)$ of piecewise constant functions for the target coefficient $c(x)$ and denote $U_h = W_h^p \times W_h^\varphi \times V_h \subset \bar{U}$. So, we consider U_h as a discrete analogue of the space $\bar{U}$. Since all norms in a finite dimensional space are equivalent, it is convenient for us to intorduce in U_h the same norm as one in $\bar{U}^1$. The functional $L(v_h)$ is defined in terms of (73) and $L'(v)(\overline{v})$ is defined in terms of (75). The FEM for (74) now reads: Find $v_h \in U_h$, such that

$$L'(v_h)(\overline{v}) = 0, \quad \forall \overline{v} \in U_h. \tag{84}$$

We solve this discrete problem using the quasi-Newton method with the limited storage [27]. More precisely, let $c_h(x) \in V_h$ be a piecewise constant approximation of the unknown coefficient $c(x)$. We compute iteratively the sequence $\{c_h^m\}$, $m = 1, ...$ of approximations of c_h as

$$c_h^{m+1}(x) = c_h^m(x) - \alpha H^m g^m(x), \tag{85}$$

where α is the step length computed via the line-search algorithm [28]. Here, H^m is the Hessian of the Lagrangian. The Hessian is computed by the usual BFGS update formula of the Hessian:

$$H^{m+1} = (I - d^m s^m y^{mT}) H^m (I - d^m y^m s^{mT}) + \rho s^m s^{mT}, m = 1, ...,$$

where

$$d^m = 1/(y^{mT} s^m), m = 1, ... \tag{86}$$

and

$$y^m = g^{m+1} - g^m.$$

Corrections s^m in (86) are defined as $s^m = c_h^{m+1} - c_h^m$. In our computations we have used a special BFGS update formula with limited storage for the Hessian [27] where we store a finite number $n = m - 1$ of corrections for the computed gradients and parameters in (86). When $n = 0$ then the quasi-Newton method is the usual gradient method with $H^0 = I$. The nodal values of the gradient $g^m(x)$ are given by (see (83))

$$g^m(x) = \gamma(c_h^m - c_0) - \int_0^T u_{ht}^m \varphi_{ht}^m \, dt. \tag{87}$$

Here $u_h^m \in W_h^u, \varphi_h^m \in W_h^u$ are functions u and φ obtained on the m^{th} iteration via solving boundary value problems (77)-(79) and (80)- (82) respectively with $c := c_h^m$, $c_0 := c_h^1$, see subsection 8.3 for our stopping criterion.

8.1 A posteriori error estimate for the Lagrangian

When performing computational experiments, we are concerned with the accuracy of obtained results. We now address the issue of a posteriori error bound that estimates the error of the finite element approximation of the function c in terms of the residual error obtained in the reconstruction process. The latter error bound can be evaluated once the FEM solution has been computed, since this solution is used then for the derivation of

that error bound. The resulting a posteriori error estimate enables us to estimate and adaptively control the finite element error to a desired tolerance level via refining the mesh locally.

Let $v \in U$ be a minimizer of the Lagrangian L on the space $\bar{U}$, and v_h be a minimizer of this functional on U_h. That is, v is a solution of the problem (75) and v_h is a solution of the problem (84). Since the second stage of our two-stage procedure, the adaptivity, is a locally convergent numerical method and the first good approximation for the second stage is obtained on the first stage, we can assume that we work in a small neighborhood of the exact solution $v^* \in U$ of our original CIP. Thus, we assume that

$$\|v - v^*\|_{\bar{U}} \leq \delta \text{ and } \|v - v_h\|_{\bar{U}} \leq \delta, \tag{88}$$

where δ is a sufficiently small positive number. We now obtain a posteriori error estimate for the error in the Lagrangian,

$$\begin{aligned} L(v) - L(v_h) &= \int_0^1 \frac{d}{d\epsilon} L(\epsilon v + (1-\epsilon)v_h) d\epsilon \\ &= \int_0^1 L'(\epsilon v + (1-\epsilon)v_h)\,(v - v_h)\,d\epsilon = L'(v_h)\,(v - v_h) + R, \end{aligned} \tag{89}$$

where R is the second order, with respect to $v - v_h$, remainder term, $|R| \leq C\,\|v - v_h\|^2_{\bar{U}}$ with a certain positive constant C (in principle, more details can be given here, which, however, is outside of the scope of this paper. We ignore R because of (88).

Let $P_h : \bar{U}^1 \to U_h$ be the operator of the orthogonal projection of the space $\bar{U}^1$ on the subspace U_h. Since $v \in U$ and $U \subset \bar{U}^1$ as a set, we can apply the operator P_h to the element v. In other words, $P_h(v) := v_h^I$ is the interpolant of v via finite elements of U_h. Using the Galerkin orthogonality (84) with the splitting $v - v_h = (v - v_h^I) + (v_h^I - v_h)$, we obtain the following error representation:

$$L(v) - L(v_h) \approx L'(v_h)\,(v - v_h^I), \tag{90}$$

involving the residual $L'(v_h)(\cdot)$ with $v - v_h^I$ appearing as the interpolation error. This splitting is one of the main tricks in the adaptivity idea, because it allows us to use the Galerkin orthogonality (84) and then to use the standard estimates of interpolation errors. We estimate $v - v_h^I$ in terms of derivatives of v and the mesh parameters h in space and τ in time. Finally we approximate the derivatives of v by corresponding derivatives of v_h, see details in [7]-[9]. It turns out that the dominating contribution of the error in the Lagrangian (72) is presented in residuals of the reconstruction and it is estimated from the above by

$$\gamma \max_{\overline{\Omega}} |c_h - c_0| + \max_{\overline{\Omega}} \int_0^T |u_{ht}\varphi_{ht}|\,dt.$$

This observation indicates that the error in the Lagrangian can be decreased by refining the grid locally in those regions, in which the absolute value of the gradient with respect to c attains its maximum. The latter forms the basis for the adaptivity technique, see Section 8.3.

8.2 A posteriori error estimate for the unknown coefficient

A more general a posteriori error estimate is the one, which can be used to estimate the error in the reconstructed coefficient rather than the error in the Lagrangian [9]. This estimate involves the solution $\widetilde{v} \in \bar{U}$ of the following problem

$$-L''(v_h)\,(\bar{v}, \widetilde{v}) = ((\psi, \bar{v})) \quad \forall \bar{v} \in U_h, \tag{91}$$

where the function $\psi \in \bar{U}$ is a function of our choice, $((\cdot,\cdot))$ is the L_2 inner product in $\bar{U}$ in space and time, and $L''(v_h)(\cdot,\cdot)$ is the second Frechet derivative (the Hessian) of the Lagrangian $L(v)$ at v_h. The second Frechet derivative of the Lagrangian expresses the sensitivity of the derivative of the Lagrangian (72) with respect to changes in v.

The main goal in the adaptive error control is to find a mesh with a few nodes as possible such that $|c - c_h| \leq \varepsilon$ for a given tolerance ε, where $c_h \in V_h$ is the third component of the vector function v_h, i.e., c_h is an approximation of the function c, which is found in our computations. Thus, in the adaptive algorithm, the solution obtained on a coarse mesh will be interpolated to the refined mesh and used then as an initial guess for a new optimization procedure. Following the ideas of [9], a posteriori estimate of the error between the exact coefficient c and the computed one c_h involves the solution of the problem (91).

Assuming existence of the solution of the problem (91), we obtain by choosing $\overline{v} = v - v_h$ in (91)

$$\begin{aligned} ((\psi, v - v_h)) &= -L''(v_h)(v - v_h, \widetilde{v}) \\ &= -L'(v)(\widetilde{v}) + L'(v_h)(\widetilde{v}) + R = L'(v_h)(\widetilde{v}) + R, \end{aligned} \tag{92}$$

where $L'(v)(v_h) = 0$ due to (74), and again R is the second order remainder term.

Since $\widetilde{v} \in \bar{U}$ and $\bar{U} \subset \bar{U}^1$ as a set, we can apply the projection operator P_h to $\widetilde{v}$. Hence, $P_h\widetilde{v} := \widetilde{v}_h^I$ is the interpolant of $\widetilde{v}$. Using splitting $\widetilde{v} = \left(\widetilde{v} - \widetilde{v}_h^I\right) + \widetilde{v}_h^I$, the Galerkin orthogonality

$$L'(v_h)(\widetilde{v}_h^I) = 0, \quad \forall \widetilde{v}_h^I \in V_h,$$

and ignoring R, we obtain from (92)

$$((\psi, v - v_h)) \approx L'(v_h)(\widetilde{v}) = L'(v_h)(\widetilde{v} - \widetilde{v}_h^I) + L'(v_h)(\widetilde{v}_h^I) = L'\left(v_h\right)(\widetilde{v} - \widetilde{v}_h^I).$$

Hence, we have obtained the following analog of a posteriori error estimate for the error in the Lagrangian (90)

$$((\psi, v - v_h)) \approx L'(v_h)(\widetilde{v} - \widetilde{v}_h^I). \tag{93}$$

We conclude, that the concrete form of the estimate (93) is the same as one for the Lagrangian $L(v)$ with only $v - v_h^I$ replaced with $\widetilde{v} - \widetilde{v}_h^I$, compare (93) with (90). Hence, to estimate $|((\psi, v - v_h))|$, we can use estimates for the derivative of the Lagrangian, thus ending up with the same problem as one in subsection 8.1.

We now provide the first rigorous explanation of the meaning of the estimate from the above of $|((\psi, v - v_h))|$ in (93), which is formulated in (94). Let $\{\psi_k\}_{k=1}^M \subset U_h$ be an orthonormal basis in the finite dimensional space U_h. Assume that for each function ψ_k there exists unique solution $\widetilde{v}_{\psi_k} \in \bar{U}$ of the problem (91) such that $\|\widetilde{v}_{\psi_k} - v^*\|_{\bar{U}} \leq \delta$ (see (88)). Denote $\widetilde{v}_{\psi_k}^I = P_h\widetilde{v}_{\psi_k}$. By (93) we have the following approximate estimate

$$|((\psi_k, v - v_h))|^2 \leq \left|L'(v_h)(\widetilde{v}_{\psi_k} - \widetilde{v}_{\psi_k}^I)\right|^2.$$

Using splitting $v - v_h = \left(v_h^I - v_h\right) + \left(v - v_h^I\right)$ again, noting that $v - v_h^I = (I - P_h)\, v$ and that by the definition of the orthogonal projection $\left(\left(\psi_k, v - v_h^I\right)\right) = ((\psi_k, (I - P_h)\, v)) = 0$, we conclude that numbers $((\psi_k, v - v_h))$ are Fourier coefficients of the vector function $v_h^I - v_h \in U_h$ with respect to the orthonormal basis $\{\psi_k\}_{k=1}^M$ in the space U_h. Hence,

$$\left[v_h^I - v_h\right]^2 = \sum_{k=1}^M |((\psi_k, v - v_h))|^2 \leq \sum_{k=1}^M \left|L'(v_h)(\widetilde{v}_{\psi_k} - \widetilde{v}_{\psi_k}^I)\right|^2,$$

$$\left[v_h^I - v_h\right] \leq \left(\sum_{k=1}^{M} \left|L'(v_h)(\widetilde{v}_{\psi_k} - \widetilde{v}^I_{\psi_k})\right|^2\right)^{1/2}. \tag{94}$$

Hence, estimates $\left|L'(v_h)(\widetilde{v}_{\psi_k} - \widetilde{v}^I_{\psi_k})\right|$ from the above for all $k = 1, ..., M$ would provide us with an estimate of the difference between the interpolant of our target minimizer of the Lagrangian and the minimizer of this Lagrangian on the subspace U_h, which will be found in computations. Note that an analogous estimate was not obtained previously in [9].

Similarly to [7]-[9] we estimate $\widetilde{v} - \widetilde{v}_h^I$ in terms of derivatives of $\widetilde{v}$ and the mesh parameters h in space and τ in time. Finally we approximate the derivatives of $\widetilde{v}$ by corresponding derivatives of $\widetilde{v}_h$, see details in [7]-[9]. Then the estimate of the right hand side of (93) is expressed in terms of residuals of the reconstruction and associated dual weights and has the following form

$$\begin{aligned}
|\left((\psi, v - v_h)\right)| &\leq \left|L'(v_h)(\widetilde{v} - \widetilde{v}_h^I)\right| \leq \int\limits_{Q_T} R_{u_1}\sigma_{\widetilde{\varphi}}\ dxdt \\
&+ \int\limits_{Q_T} R_{u_2}\sigma_{\widetilde{\varphi}}\ dxdt + \int\limits_{S_T} R_{\varphi_1}\sigma_{\widetilde{u}}\ dxdt \\
&+ \int\limits_{Q_T} R_{\varphi_2}\sigma_{\widetilde{u}}\ dxdt + \int\limits_{Q_T} R_{\varphi_3}\sigma_{\widetilde{u}}\ dxdt \\
&+ \int\limits_{Q_T} R_{c_1}\sigma_{\widetilde{c}}\ dx + \int_{\Omega} R_{c_2}\sigma_{\widetilde{c}}\ dxdt,
\end{aligned} \tag{95}$$

where $\widetilde{v} = (\widetilde{u}, \widetilde{\varphi}, \widetilde{c})$ is a solution of the problem (91) for a chosen function $\psi \in U_h$ and $S_T = \partial\Omega \times (0, T)$. Residuals are defined as

$$\begin{aligned}
R_{u_1} &= \max_{S\subset\partial K} h_k^{-1}\left|\left[\partial_s u_h\right]\right|, \quad R_{u_2} = c_h\tau^{-1}\left|\left[\partial_t u_h\right]\right|, \\
R_{\varphi_1} &= \left|u_h|_{S_T} - g\right|, R_{\varphi_2} = \max_{S\subset\partial K} h_k^{-1}\left|\left[\partial_s \varphi_h\right]\right|, \quad R_{\varphi_3} = c_h\tau^{-1}\left|\left[\partial_t \varphi_h\right]\right|, \\
R_{c_1} &= \left|\frac{\partial\varphi_h}{\partial t}\right| \cdot \left|\frac{\partial u_h}{\partial t}\right|, R_{c_2} = \gamma|(c_h - c_0)|
\end{aligned} \tag{96}$$

and different weights σ have the following form:

$$\begin{aligned}
\sigma_{\widetilde{\varphi}} &= C_1\tau\left|[\partial_t\widetilde{\varphi}_h]\right| + C_1 h\left|[\partial_s\widetilde{\varphi}_h]\right|, \\
\sigma_{\widetilde{u}} &= C_1\tau\left|[\partial_t\widetilde{u}_h]\right| + C_1 h\left|[\partial_s\widetilde{u}_h]\right|, \\
\sigma_{\widetilde{c}} &= C_2\left|[\widetilde{c}_h]\right|,
\end{aligned} \tag{97}$$

where $[\widetilde{v}]$ on a space element K (or time-interval J) denotes the maximum of the modulus of the jump of the quantity $\widetilde{v}$ across a face of K (or boundary node of J). In particular $[\partial_s\widetilde{v}]$ on a space-element K denotes the maximum of the modulus of the jump in the normal derivative of $\widetilde{v}$ across a side of K. Also, $[\partial_t\widetilde{v}]$ on a time-interval J is the maximum of the modulus of the jump of the time derivative of $\widetilde{v}$ across a boundary node of J. Here C_1 and C_2 are interpolation constants.

Thus, to find weights (97) in estimates (95), we need to compute the problem (91) to find the function $\widetilde{v}$. It follows from (94) that choosing different functions ψ_k from the orthonormal basis $\{\psi_k\}_{k=1}^M$ of the subspace U_h in the problem (91), we obtain an approximate a posteriori control of the error between the interpolant v_h^I of the exact minimizer

$v \in U$ and computed minimizer $v_h \in U_h$ of the Lagrangian. The main difficulty here is in the solvability of the problem (91). A certain numerical method for the solution of this problem was proposed in [9], and it was confirmed by numerical examples. However, questions of convergence and stability of that method, so as the question of existence of solution of the problem (91) are open. Still, we can come up with a simplified estimate (94) which does not require solution of the problem (91), although assumes its existence, see next subsection. Our computational experience shows that this estimate is sufficient for our goal.

8.3 The adaptive algorithm

In this section we present our adaptive algorithm based on computations of the residuals for the computed coefficient c. The initial guess value c_0 for the unknown coefficient c in our adaptive algorithm on the initial mesh is taken from the solution obtained by the above globally convergent algorithm.

One can see from (94) and (95) that the error in the reconstructed coefficient consists out of a sum of integrals of different residuals multiplied by the interpolation errors. Thus, to estimate the error in the unknown coefficient, we need to compute the approximated values of (u_h, φ_h, c_h) together with the residuals and interpolation errors. We refine the mesh adaptively at the end of the optimization procedure (85), (87). Hence, we can assume that the values of the residuals $R_{\varphi_i}, i = 2, 3$ and interpolation errors $\sigma_{\widetilde{\varphi}}$ for the adjoint solution φ_h are small and we ignore them in (95). Value of the residual R_{φ_1} is small because $||u_h|_{S_T} - g|| << 1$ and we can ignore it as well in (95). Thus, we compute only dominating residuals R_{c_1} and R_{c_2}. Our computational experience shows that this is enough, i.e., that the approximate error estimate (98) is sufficient for the solution enhancement via the adaptivity technique.

It follows from (94) that if a solution of the problem (91) exists for each function ψ_k, then we can write the following approximate estimate for the error in the computed unknown coefficient

$$\left\| c_h^I - c_h \right\|_{L_2(\Omega)} \leq MC_2 A(\Omega) \left| [\widetilde{c}_h] \right| \int\limits_0^T \left(\max_{\overline{\Omega}} R_{c_1}(x,t) + \max_{\overline{\Omega}} R_{c_2}(x,t) \right) dt, \tag{98}$$

where $A(\Omega)$ is the area of the domain Ω (volume in the 3-d case) and M is the dimension of the subspace U_h. If, however, solution of the problem (91) does not exist for some (or all) functions ψ_k, then it follows from (90) that the integral term in (98) estimates from the above the error in the Lagrangian,

$$\left| L(v) - L(v_h) \right| \approx \left| L'(v_h)\left(v - v_h^I\right) \right| \leq A(\Omega) \left| [c_h] \right| \int\limits_0^T \left(\max_{\overline{\Omega}} R_{c_1}(x,t) + \max_{\overline{\Omega}} R_{c_2}(x,t) \right) dt. \tag{99}$$

Thus, we can hope to decrease the error via locally refining mesh in those regions, where values of residuals $R_{c_1}(x,t), R_{c_2}(x,t)$ are close to the maximal ones. Estimates (**??**) and (95) allow us to control the error in the computed reconstructed coefficient c_h. Since residuals R_{c_1} and R_{c_2} are independent on the solution of the problem (91), our algorithm does not use that solution. Although estimates (98) and (99) are approximate ones, our computational experience shows that they are sufficient.

In our computations we use the following version of the adaptive algorithm.

0. Choose an initial mesh K_h and an initial time partition J_0 of the time interval $(0,T]$. Start with an initial guess $c_0 = c_{glob}$, which was computed in the above globally convergent algorithm, and compute the sequence of c^m in the following steps:

1. Compute the solution u^m of the forward problem (77)-(79) on K_h and J_k, with $c(x) = c^m$.

2. Compute the solution φ^m of the adjoint problem (80)-(82) backwards in time on K_h and J_k.

3. Update the coefficient $c := c_h$ on K_h and J_k using the quasi-Newton method (85)
$$c^{m+1} = c^m - \alpha H^m g^m.$$

4. Stop computing c if either the norm of the gradient g^m of the Lagrangian with respect to the coefficient in (87) is $||g^m||_{L_2(\Omega)} < \theta$ or norms $||g^n||_{L_2(\Omega)}$ are stabilized. Otherwise set $m = m+1$ and go to step 1. Here, θ is the tolerance in quasi-Newton updates. In our computations we took $\theta = 10^{-5}$.

5. Compute the residuals, R_{c_1}, R_{c_2} and refine the mesh at all points where
$$\int_0^T \left(\max_{\overline{\Omega}} R_{c_1}(x,t) + \max_{\overline{\Omega}} R_{c_2}(x,t) \right) dt > tol. \tag{100}$$
Here *tol* is a tolerance chosen by the user.

6. Construct a new mesh K_h and a new time partition J_k. On J_k the new time step τ should be chosen with respect to the CFL condition. Interpolate the reconstructed coefficient c_h from the previous mesh to the new mesh. Return to the step 1 and perform all the steps of the optimization algorithm on the new mesh.

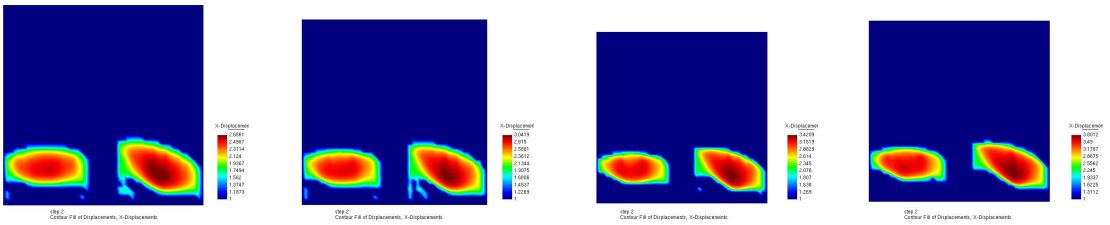
9 Numerical Testing

9.1 Results of reconstruction using the globally convergent algorithm

We have performed numerical experiments to reconstruct the medium, which is homogeneous with $c(x) = 1$ except of two small squares, where $c(x) = 4$, see Figure 1-c). However, we have not assumed *a priori* knowledge of neither the structure of this medium nor of the background constant $c(x) = 1$ for $x \in \Omega \setminus$ those two squares, although, following the Tikhonov concept (as mentioned in section 2), we have assumed the knowledge of the constant $d_1 = 1/2$, see (3) and (67). Because of this, the starting value for the tail $V_{1,1}(x,\overline{s})$ was computed via solving the forward problem (68) for $c \equiv 1$. Let $w_{c\equiv 1}(x,\overline{s})$ be the corresponding function $w(x,s)$ at $s = \overline{s}$. Then, using (18), we took $V_{1,1}(x,\overline{s}) = \overline{s}^{-2} \ln w_{c\equiv 1}(x,\overline{s})$.

It was found in [5] that for domains G, Ω specified in section 7 the interval $[\underline{s}, \overline{s}] = [6.7, 7.45]$ is the optimal one, and so we have used it in our numerical studies. We have chosen the step size with respect to the pseudo frequency $h = 0.05$. Hence, $N = 15$ in our case. We have chosen two sequences of regularization parameters $\lambda := \lambda_n$ and $\varepsilon = \varepsilon_n$ for $n = 1, ..., \overline{N}$, which are the same as ones in [5],

$$\lambda_n = 20, n = 1,2; \lambda_n = 200, n = 3,4,5; \lambda_n = 2000, n \geq 6;$$
$$\varepsilon_n = 0, n = 1,2; \varepsilon_n = 0.001, n = 3,4,5; \varepsilon_n = 0.01, n = 6,7,$$
$$\varepsilon_n = 0.1, n \geq 8.$$



a) $c_{9,2}$ b) $c_{10,2}$ c) $c_{11,2}$ d) $c_{12,2}$

Figure 4: Test 1.1: spatial distribution of c_h after computing $q_{n,k}; n = 9, 10, 11, 12$, where n is number of the computed function q.

Once the function q_n is calculated, we update the function $c := c_n$ as in (28), (29), see subsection 7.3 of [5] for some numerical details. The resulting computed function is $c(x) := c_{\overline{N}}(x)$. Comparing with [5], in the current work we choose a completely different stopping rule. In calculating iterations with respect to the nonlinear term (Section 5), we consider norms F_n^k,

$$F_n^k = ||q_{n,1}^k|_{\partial\Omega} - \overline{\psi}_n||_{L_2(\partial\Omega)}.$$

We stop our iterations with respect to nonlinear terms when either

$$\text{either } F_n^k \geq F_n^{k-1} \text{ or } F_n^k \leq \varepsilon,$$

where $\varepsilon = 0.001$ is a small tolerance number of our choice. In other words, we stop iterations, when either F_n^k starts to grow or are too small. Next, we iterate with respect to the tails and use the same stopping criterion. Namely, we stop our iterations with respect to tails when either

$$F_{n,i} \geq F_{n,i-1} \tag{101}$$

or

$$F_{n,i} \leq \varepsilon, \tag{102}$$

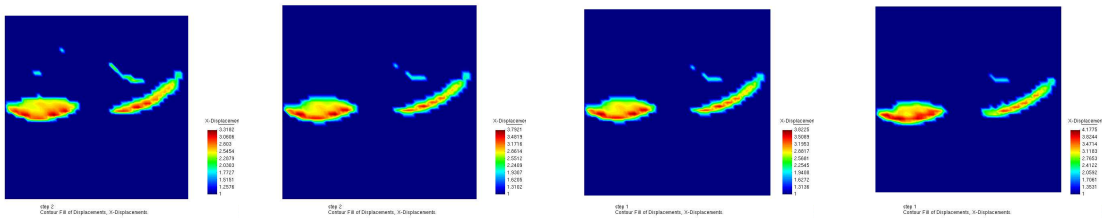
where $F_{n,i} = ||q_{n,i}|_{\partial\Omega} - \overline{\psi}_n||_{L_2(\partial\Omega)}$. So, in accordance with Section 5 the number i, on which these iterations are stopped, is denoted as $i := m_n$. Once the criterion (101)-(102) is satisfied, we take the last computed tail V_{n,m_n}, set $V_{n+1,1} := V_{n,m_n}$ and run computations again. This difference allows us to get a more flexible stopping rule in global convergence algorithm. Hence, the number m_n of iterations with respect to tails is chosen automatically "inside" of each iteration for q_n. Thus, numbers m_n vary with n. This is different from [5], where numbers m_n where not chosen automatically.

In all our tests we have introduced the multiplicative random noise in the boundary data, g_σ, by adding relative error to computed data g using the following expression

$$g_\sigma\left(x^i, t^j\right) = g\left(x^i, t^j\right)\left[1 + \frac{\alpha_j(g_{max} - g_{min})\sigma}{100}\right].$$

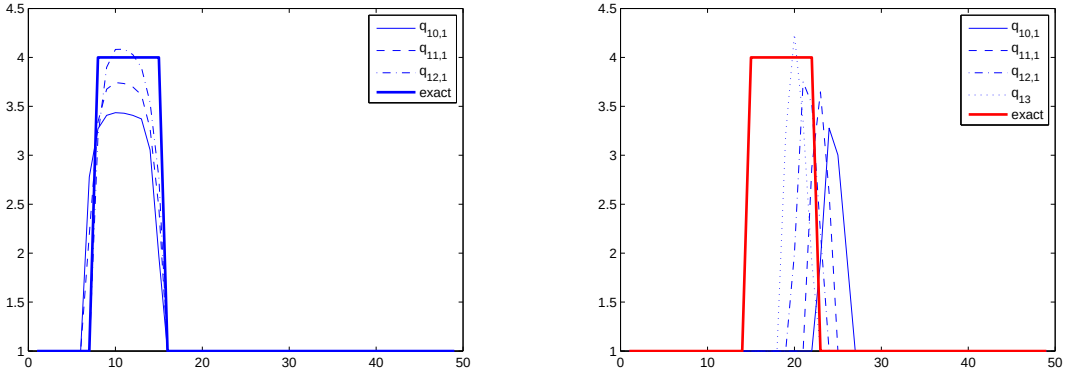
Here, $g\left(x^i, t^j\right) = u\left(x^i, t^j\right), x^i \in \partial\Omega$ is a mesh point at the boundary $\partial\Omega, t^j \in (0, T)$ is a mesh point in time, α_j is a random number in the interval $[-1; 1]$, g_{max} and g_{min} are maximal and minimal values of the computed data g, respectively, and $\sigma = 5\%$ is the noise level.

Computations were performed on 16 parallel processors in NOTUR 2 production system at NTNU, Trondheim, Norway (67 IBM p575+ 16-way nodes, 1.9GHz dual-core CPU, 2464 GB memory).



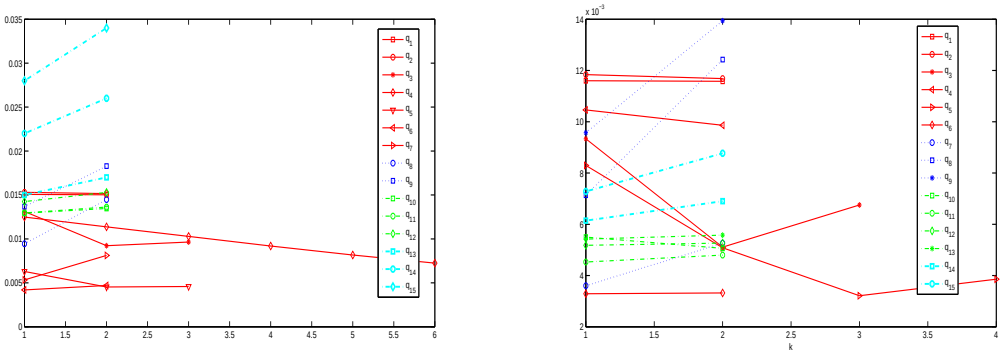
a) $c_{10,2}$ b) $c_{11,1}$ c) $c_{12,1}$ d) $c_{13,1}$

Figure 5: Test 2.1: spatial distribution of c_h after computing $q_{n,k}; n = 10, 11, 12, 13$ where n is number of the computed function q.



a) Test 1.1 b) Test 2.1

Figure 6: The one-dimensional cross-sections of the image of the function $c_{n,k}$ computed for corresponding functions $q_{n,1}$. On a) for Test 1.1 along the vertical line passing through the middle of the right small square; and on b) for Test 2.1 along the vertical line passing through the middle of the left small square.



a) Test 1.1 b) Test 2.1

Figure 7: Computed L_2-norms of the $F_{n,i} = ||q_{n,i} \mid_{\partial\Omega} - \overline{\psi}_n||_{L_2(\partial\Omega)}$.

Test 1.1

We test our numerical method on the reconstruction of the structure given on Figure 1-c). The plane wave f is initialized at the top boundary ∂G_1 of the computational domain G, propagates during the time period $(0, t_1]$ into G, is absorbed at the bottom boundary ∂G_2 for all times $t \in (0, T)$ and it is also absorbed at the top boundary ∂G_1 for times $t \in (t_1, T)$, see Figures 2.

Figure 4 displays isosurfaces of resulting images of functions $c_{n,k}, n = 9, 10, 11, 12$. Figure 6-a) presents the one-dimensional cross-sections of computed images of functions $c_{n,k}$ superimposed with the correct one along the vertical line passing through the middle of the right small square. Comparison of images of functions $c_{n,k}$ for different values n and k shows that the inclusion/background contrasts grow with the grow of n and k.

One can see from Figure 4 that the $3.8 : 1$ contrast in the right square is imaged for $n := \overline{N} = 12$ (see below for this choice of $\overline{N}$). Thus, we have obtained the 5% error $(0.2/4)$ in the imaged contrast, which is exactly the same as the noise level in the data. As to the left square, we got the same contrast. However, location of the left square is shifted downwards, and both imaged squares are on about the same horizontal level. Values of the function $c(x) = 1$ outside of these squares are also imaged accurately.

Using Figure 7-a) which shows computed L_2-norms $F_{n,i}$, we analyze results of the reconstruction. One can see on Figure 7 that the number m_n of iterations with respect to tails indeed varies with n, since m_n is chosen automatically now, using the criterion (101)-(102). We observe that the computed $F_{n,i}$ decrease until computing the function q_8. Next, $F_{9,2} > F_{8,2},\ F_{10,2} < F_{9,2}$ and then these norms stabilize on $n = 11, 12$. For $n = 13, 14, 15$ norms $F_{n,2}$ grow steeply. Thus, we conclude, that $\overline{N} = 12$ and we take $c_{12,2}$ as our final reconstruction result.

Test 2.1

We now test our globally convergent method on the structure given on Figure 1-c). However, the difference with the previous test is that we use the plane wave, which is initialized at the bottom boundary of computational domain G, see Fig.2. Figure 5 displays isosurfaces of resulting images of functions $c_{n,k}, n = 10, 11, 12, 13$. Figure 6-b) displays the one-dimensional cross-sections of computed images of functions $c_{n,k}$ superimposed with the correct one along the vertical line passing through the middle of the left small square. One can see from Figure 5 that the $3.8 : 1$ contrast for $n := \overline{N} = 12$ (see below for this choice of $\overline{N}$) in the left square is imaged again with 5% error $(0.2/4)$ which is the same as the noise level in the data. As to the right square, we got the same $3.8 : 1$ contrast. However, again both squares are imaged on about the same vertical level.

Using Figure 7-b), which shows computed L_2-norms $F_{n,i}$, we analyze results of the reconstruction. We observe that computed norms $F_{n,i}$ decrease with n until computing the function q_7, i.e.., until $n = 7$ and these numbers grow with the increase of $n = 8, 9$. Next, we observe a steep decrease at $n = 10$ and a stabilization for $n = 11, 12$. For $n = 13, 14, 15$ norms $F_{n,i}$ grow steeply. Thus, we conclude, that $\overline{N} = 12$ and we take $c_{12,1}$ as our final reconstruction result. We observe, that in both Tests 1.1 and 2.1 the location of the square, which is located closer to the side from which the plane wave is launched, is imaged better, while the inclusion/background contrast is imaged well in both small squares, so as the value of the coefficient $c(x) = 1$ outside of (imaged) small squares. Thus, we are prompted to use the adaptivity technique in order to enhance images of locations.

9.2 The synthesis of the globally convergent algorithm with the adaptivity

The goal of two tests of this subsection is to demonstrate the performance of the synthesis of our globally convergence algorithm with the adaptivity technique. Since the adaptivity

is a locally convergent numerical method, we take the starting point for the adaptivity the image obtained by the globally convergent method. Below "Test 1.2" (respectively "Test 2.2") means that we take the image obtained in the above Test 1.1 (respectively in Test 1.2), as the starting point for our finite element adaptive algorithm. The boundary data $g = u\mid_{\partial\Omega}$ in both Tests 1.2 and 2.2 are the same as ones in Test 1.1 and 2.1 respectively, except that in Test 1.2 we use two noise levels in two "sub-tests": 0% and 5%. In Tests 1.2 and 2.2 let Γ be the side of the square Ω, opposite to the side from which the plane wave is launched and $\Gamma_T = \Gamma \times (0,T)$. In some sense the side Γ_T is the most sensitive one to the resulting data.

The adaptive algorithm means, that we find the solution of our problem in an iterative process, where we start with a coarse mesh shown on Figure 8-a), c), find an approximate solution by the quasi-Newton method on this mesh, see Section 8.3. Next, we evaluate residuals as in (100). Then we refine the mesh locally at those regions where residuals have largest values, construct a new mesh and a new time partition, and repeat the computations again on this new mesh. We stop iterative process when L_2-norms of the computed gradient for the coefficient are stabilized or started to increase for all further refinements of the mesh. Let $|R_c(x)| = |R_{c_1}(x)| + |R_{c_2}(x)|$, see (95) and (99). We refine the mesh at all regions where

$$|R_c(x)| \geq \beta \max |R_c(x)|, \tag{103}$$

where $\beta = const \in (0,1)$ is the tolerance number of our choice. The choice of the parameter β depends on the behavior of the computed value of $\max |R_c(x)|$ in right hand side of (103). If we take β too small (for example, $\beta = 0$), then we will refine mesh almost in the entire domain Ω, since, realistically, after the optimization procedure $|R_c(x)|$ will be non- zero at almost all mesh points. Unlike this, our goal is to construct a new mesh with a few nodes as possible, while still getting a good enhancement of the solution obtained on the globally convergent stage of our two-stage numerical procedure. Hence, we take only maximal values of the computed residual $|R_c(x)|$ and refine mesh in a small neighborhood of those points where this maximal value is achieved. On the other hand, the parameter β can not be taken too close to 1 also, since in this case the automatic adaptive algorithm will come up with a too narrow region, where the mesh should be refined. Thus, the choice of β depends on concrete values of the gradient function $|R_c(x)|$ and should be chosen in numerical experiments. In (103) we take $\beta = 0.1$ on the coarse mesh, $\beta = 0.2$ on the one, two and three refined meshes, and $\beta = 0.6$ for all next refinements of the initial mesh.

On all refined meshes we have used a cut-off parameter C_{cut} for the reconstructed coefficient c_h such that

$$c_h = \begin{cases} c_h, & \text{if } |c_h - c_{guess}| \geq C_{glob} \\ c_{glob}, & \text{elsewhere.} \end{cases}$$

We choose $C_{cut} = 0$ for $m < 3$, $C_{cut} = 0.3$ for $m \geq 3$ in all tests. Here, m is the number of iterations in quasi-Newton method. Hence, the cut-off parameter ensures that we do not go too far from c_{glob}. The application of the adaptivity technique allows us to get more correct locations of both small squares depicted in Figure 1-c).

In the adaptive algorithm we can use box constrains for the reconstructed coefficient. We obtain these constraints using the solution obtained in the globally convergent part. Namely, in Tests 1.2 and 2.2 minimal and maximal values of the target coefficient in box constraints are taken using results of Tests 1.1 and 2.1. So, when conducting Tests 1.1 and 2.1, we have used only the knowledge of the number $d_1 = 0.5$ in (3). Now, since we know that the solution obtained on the first stage is a good approximation for the correct solution (Theorem 6.1) and the maximal value of the computed coefficient is 3.8, we set

opt.it.	4608 elements	5340 elements	8230 elements	14604 elements	23344 elements
1	0.0193568	0.0167242	0.0146001	0.0131787	0.0224184
2	0.0193944	0.0157746	0.0139716	0.0133006	0.0208246
3		0.0133565			0.0208889
4		0.0125237			0.0204343

Table 1: Test 1.2: $||u\mid_{\Gamma_T} - g||_{L_2(\Gamma_T)}$ on adaptively refined meshes. The number of stored corrections in quasi-Newton method is $n = 15$. Computations was performed with the noise level $\sigma = 0\%$ and with the regularization parameter $\gamma = 0.01$.

opt.it.	4608 elements	5340 elements	6356 elements	10058 elements	14586 elements
1	0.0992683	0.097325	0.0961796	0.0866793	0.0880115
2	0.0988798	0.097322	0.096723	0.0868341	0.0880866
3	0.0959911	0.096723			0.0876543
4		0.096658			

Table 2: Test 1.2: $||u\mid_{\Gamma_T} - g||_{L_2(\Gamma_T)}$ on adaptively refined meshes. The number of stored corrections in quasi-Newton method is $n = 15$. Computations was performed with the noise level $\sigma = 5\%$ and with the regularization parameter $\gamma = 0.01$.

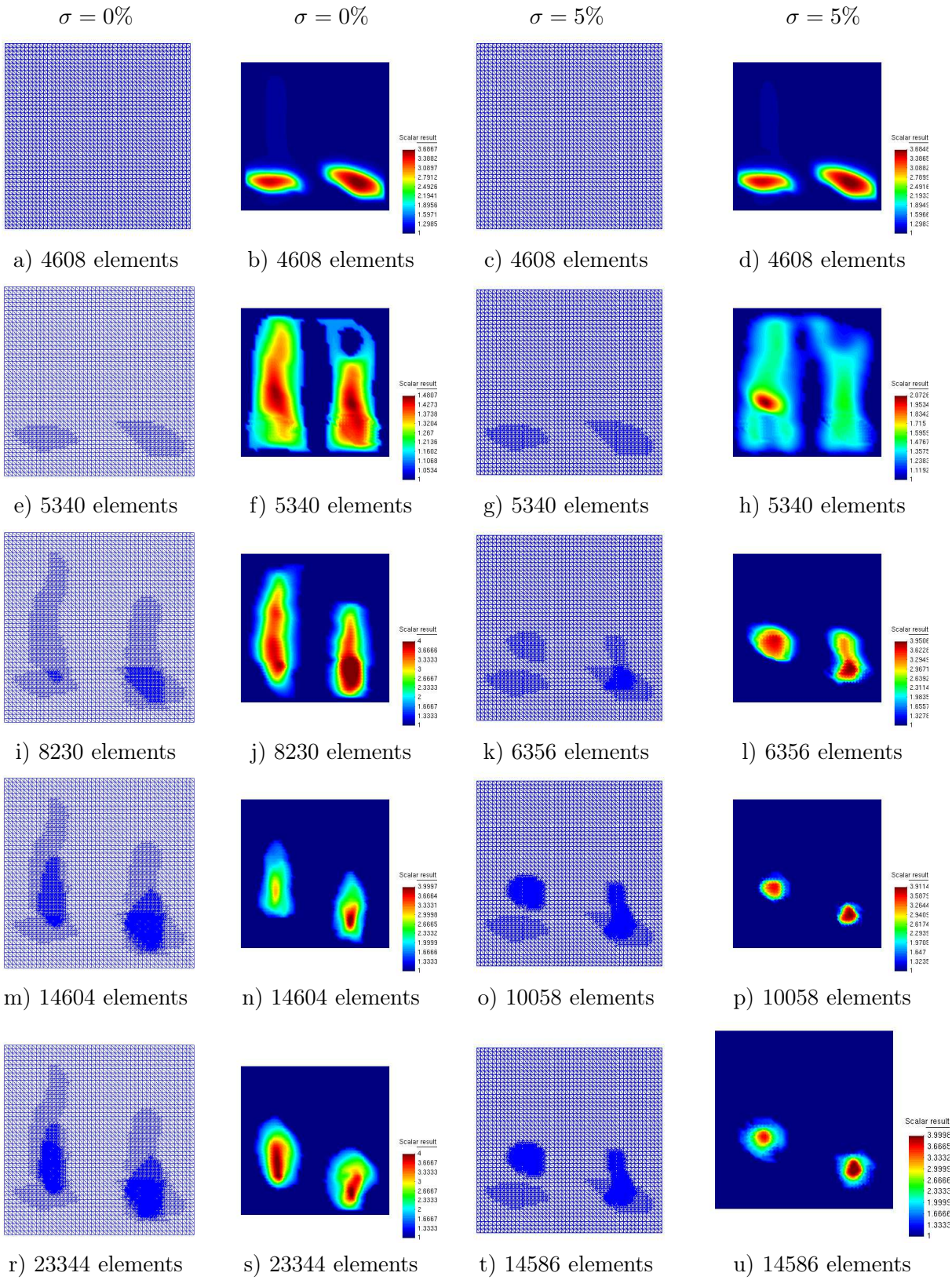
$d_2 = 2$ in (3). Thus, in tests 1.2 and 2.2 we enforce that the coefficient $c(x)$ belongs to the set of admissible parameters, $c(x) \in C_M = \{c \in C(\overline{\Omega}) | 1 \leq c(x) \leq 4\}$.

Test 1.2.

We test the synthesis of both globally convergent and adaptive methods with the starting point on the coarse mesh taken from the results of Test 1.1 and with the plane wave initialized at the top boundary of the computational domain G. More precisely, as the starting point for the coefficient $c(x)$ in the adaptive algorithm on the coarse mesh we take $c_{12,2}$, which corresponds to Figure 4-d). The initial coarse mesh is shown on Figures 8-a,c). We have performed two set of numerical experiments: with introducing $\sigma = 0\%$ and $\sigma = 5\%$ of the multiplicative random noise in the function $g(x,t)$ in an adaptive procedure. Testing was performed on 4 times adaptively refined meshes shown on Figure 8. We note that in both Tests 1.2 and 2.2 boundary points are the same for all refinements of the initial mesh, since they are located at the common boundary with the subdomain G_{FDM} (Figure 1) and should be kept unchanged in order to perform the exchange procedure in the hybrid method when solving the forward problem (68). Figure 8 shows that the adaptivity technique enhances the quality of the reconstruction obtained on the first stage. We are able to reconstruct well locations of both small squares while preserving a good initially obtained inclusion/background contrast, which turns out to be now 4:1 instead of 3.8:1 calculated on the first stage. The value of the coefficient $c(x) = 1$ outside of small squares is also imaged well. We observe that the use of the initial coarse mesh with 4608 elements does not improve the image obtained on Test 1.1.

9.2.1 The case $\sigma = 0\%$

Images 8-b), f), j), n), s) were obtained with $\sigma = 0\%$ in the boundary data g, with the regularization parameter $\gamma = 0.01$ and without using the smoothness indicator procedure



a) 4608 elements b) 4608 elements c) 4608 elements d) 4608 elements

e) 5340 elements f) 5340 elements g) 5340 elements h) 5340 elements

i) 8230 elements j) 8230 elements k) 6356 elements l) 6356 elements

m) 14604 elements n) 14604 elements o) 10058 elements p) 10058 elements

r) 23344 elements s) 23344 elements t) 14586 elements u) 14586 elements

Figure 8: Test 1.2: Adaptively refined computational meshes: with $\sigma = 0\%$ - on a),e),i),m),r), with $\sigma = 5\%$ - on c),g),k),o),t), and correspondingly spatial distribution of the parameter c_h: with $\sigma = 0\%$ - on b),f),j),n),s) and with $\sigma = 5\%$ - on d),h),l),p),u).

opt.it.	4608 elements	5298 elements	7810 elements	11528 elements	19182 elements
1	0.0992683	0.0976474	0.0976851	0.089979	0.0977153
2	0.0988798		0.0974385	0.0901018	0.097487
3	0.0959911			0.0901153	0.0975039

Table 3: Test 2.2: $||u\mid_{\Gamma_T} - g||_{L_2(\Gamma_T)}$ on adaptively refined meshes. The number of stored corrections in quasi-Newton method is $n = 15$. Computations was performed with the noise level $\sigma = 5\%$ and with the regularization parameter $\gamma = 0.01$.

applied to the reconstructed coefficient $c(x)$. As it was stated in subsection 7.3 of [5], this procedure consists in a local averaging of computed values of $c_{n,i}(x)$. The effect of using smoothness indicator procedure can be seen from comparison of $\sigma = 0\%$ and $\sigma = 5\%$ images of Figure 8. So, when this procedure was not applied, we got more elements in new adaptively refined meshes and more "washed away" images compared with images for the case of $\sigma = 5\%$ when this procedure was in place.

In Table 1 we present computed L_2-norms of $||u\mid_{\Gamma_T} - g||_{L_2(\Gamma_T)}$ in the quasi-Newton method for adaptively refined meshes. Here functions $u\mid_{\Gamma_T}$ are computed via the forward problem solutions and the same in two more tests below. We observe that these norms decrease as meshes are refined. Then they slightly increase and are finally stabilized for all refinements $n > 3$ of the initial mesh.

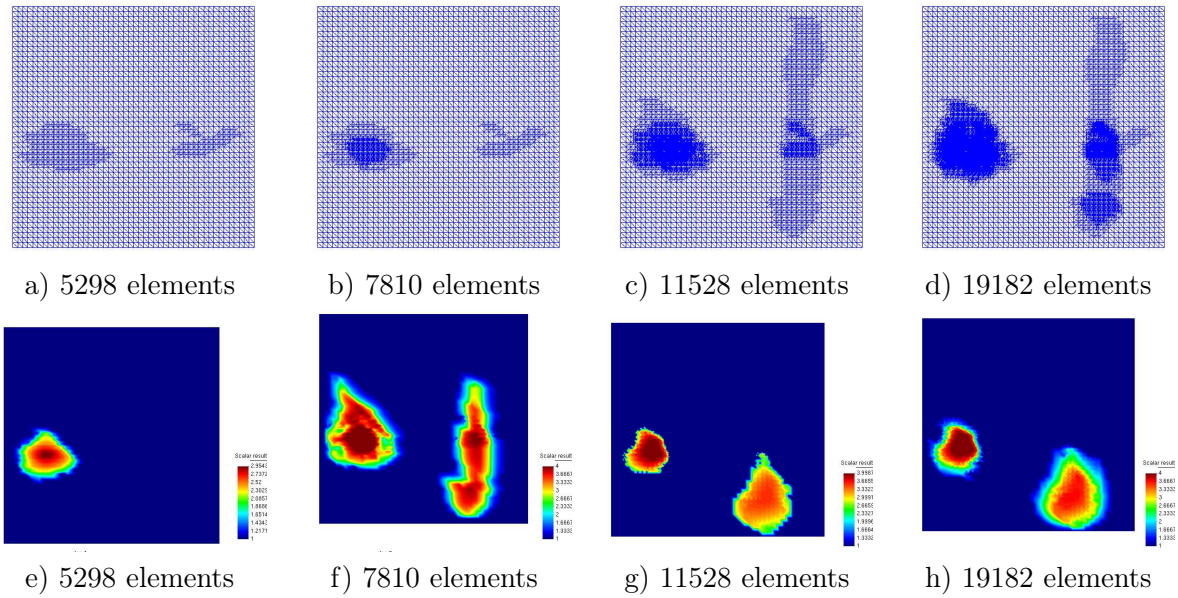
9.2.2 The case $\sigma = 5\%$

Images 8-d), h), l), p), u) were obtained with $\sigma = 5\%$ in the boundary data g, with regularization parameter $\gamma = 0.01$ and with the smoothing indicator procedure on the all adaptively refined meshes. The use of the smoothing indicator for the reconstructed coefficient c_h has helped us to obtain more accurate images as well as to get a lesser number of finite elements in computational meshes. Table 2 presents computed L_2-norms of $||u\mid_{\Gamma_T} - g||_{L_2(\Gamma_T)}$. We observe that norms at the boundary are decreasing as meshes are refined. Then they slightly increase and are finally stabilized for all refinements $n > 3$ of the initial mesh.

Test 2.2

Now we test the synthesis of the globally convergent numerical method with the adaptivity with the starting point on the coarse mesh taken from the result of Test 2.1 and with the plane wave initialized at the bottom boundary of the computational domain G. The initial guess for the adaptive algorithm on the coarse mesh is the computed coefficient $c_{12,1}(x)$ presented on Figure 5-c). The boundary data g is taken the same as in Test 2.1, i.e., with the $\sigma = 5\%$ of the multiplicative random noise.

Again, we have performed tests on 4 times adaptively refined meshes shown on Figure 9-a)-d). Just as in Test 1.2, we observe that the adaptivity technique improves the quality of the reconstruction, see reconstruction results on Figure 9-e)-h). In Table 3 we present computed norms of $||u\mid_{\Gamma_T} - g||_{L_2(\Gamma_T)}$. We observe that these norms decrease as meshes are refined. They decrease until the third refinement. On the fourth refinement they slightly increase and then they stabilize. Further mesh refinements are not necessary since norms $||u\mid_{S_T} - g||_{L_2(S_T)}$ are stabilized for all refinements with $n > 3$ of the initial mesh, and we get the same reconstruction result with further refinements. Thus, using Table 3, we conclude that on three times refined mesh we get solution of our inverse problem.



a) 5298 elements b) 7810 elements c) 11528 elements d) 19182 elements

e) 5298 elements f) 7810 elements g) 11528 elements h) 19182 elements

Figure 9: Test 2.2: Adaptively refined computational meshes on a)-d) and spatial distribution of the parameter c_h with $\sigma = 5\%$, which corresponds to these meshes, on e)-h).

10 Summary

We have presented a modified globally convergent numerical method of [5] for a multidimensional CIP for a hyperbolic PDE. As it follows from the global convergence Theorem 6.1, the globally convergent numerical method provides a good starting point for the finite element adaptive method. This naturally leads to a two-stage numerical procedure, which synthesizes both approaches. On the first stage the globally convergent numerical method is used. On the second stage solution obtained on the first is used as the starting point for the locally convergent adaptivity technique. This technique enhances the solution obtained on the first stage. An important observation of our numerical testing is that the first step of the adaptivity, when the quasi-Newton method applied on the same coarse mesh, on which the globally convergent part was working, does not provide a noticeable change for the image obtained on the globally convergent stage, see Figure 8. Hence, the use of locally refined meshes, which is the central point of to the adaptivity, is essential here.

The adaptivity is based on a posteriori analysis of: (1) the error in the Lagrangian and (2) the error in the solution. As a result, one can locate spots where the maximum error of the reconstructed coefficient likely is. Next, the spatial mesh is refined locally with the feedback from a posteriori error estimator. The main achievement of the adaptivity is that one does not need to know in advance the solution of a corresponding CIP for that a posteriori error analysis. Another new element of this work is that we have rigorously explained the meaning of the so-called ψ function in the procedure of estimating the error in the computed coefficient, which was not explained previously. Numerical tests have shown a good performance of our two-stage procedure.

Acknowledgement. The authors are grateful to M.Yu. Kokurin for fruitful discussions. This work was supported by the U.S. Army Research Laboratory and U.S. Army Research Office under grants number W911NF-05-1-0378 and W911NF-08-1-0470. The work of the first author was also supported by the Project No. IKT 820.94.000 at NTNU, Department of Mathematical Sciences, Trondheim, Norway. NOTUR 2 production system at NTNU, Trondheim, Norway is acknowledged.

References

[1] H. Ammari, E. Iakovleva, and D. Lesselier. Music-type electromagnetic imaging of a collection of small three dimensional inclusions. *SIAM J.Sci.Comp.*, 29:674 709, 2007.

[2] S. Arridge, Optical tomography in medical imaging, *Inverse Problems*, 15, 841–893, 1999.

[3] A.B. Bakushinsky and M.Yu. Kokurin, *Iterative Methods for Approximate Solution of Inverse Problems*, Springer, 2005.

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

[5] L. Beilina and M. V. Klibanov. A globally convergent numerical method for a coefficient inverse problem, *SIAM J. Sci. Comp.*, 31(1):478-509, 2008.

[6] R. Becker and R. Rannacher, An optimal control approach to a posteriori error estimation in finite element methods, *Acta Numerica*, Cambridge University Press, 1–225, 2001

[7] 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(1):382 402, 2006.

[8] 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.

[9] L. Beilina and C. Johnson. A posteriori error estimation in computational inverse scattering. *Mathematical models and methods in applied sciences*, 15(1):23 37, 2005.

[10] L. Beilina, M.Hatlo, and H. Krogstad. Adaptive error control in inverse electromagnetic scattering. *Technical Report* No. 4, NTNU, Norway, 2008. Submitted for publication, available online at http://www.math.ntnu.no/preprint/numerics/.

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

[12] M. I. Belishev. Boundary control in reconstruction of manifolds and metrics (the bc method). *Inverse Problems*, 13, R1-R45, 1997.

[13] M. I. Belishev and V. Yu Gotlib. Dynamical variant of the bc-method: theory and numerical testing. *Inverse and Ill-Posed Problems*, 7:221 240, 1999.

[14] V. A. Burov, S. A. Morozov, and O. D. Rumyantseva. Reconstruction of ne-scale structure of acoustical scatterers on large-scale contrast background. *Acoustical Imaging*, 26:231 238, 2002.

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

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

[17] B. Engquist and A. Majda. Absorbing boundary conditions for the numerical simulation of waves. *Math. Comp.*, 31:629 651, 1977.

[18] Yu.A. Gryazin, M.V. Klibanov and T.R. Lucas, Numerical solution of a subsurface imaging inverse problem, SIAM J. Appl. Math., 62: 664-683, 2001.

[19] M. V. Klibanov and A. Timonov. *Carleman Estimates for Coefficient Inverse Problems and Numerical Appli- cations.* VSP, Utrecht, The Netherlands, 2004

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

[21] M. Klibanov and A. A. Timonov, Numerical studies on the globally convergent convexification algorithm in 2D, *Inverse Problems*, 23, 123–138, 2007.

[22] O. A. Ladyzhenskaya and N. N. Uralceva, *Linear and Quasilinear Elliptic Equations*, Academic Press, New York, 1969

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

[24] J. Mueller and S. Siltanen. Direct reconstructions of conductivities from boundary measurements. *SIAM J. Sci. Comp.*, 24:1232 1266, 2003.

[25] R. G. Novikov. Multidimensional inverse spectral problem for the equation $-\triangle\psi + (v(x) - Eu(x))\psi = 0$. *Functional Analysis and Its Applications*, 22:11 22, 1988.

[26] R. G. Novikov. The $\overline{\partial}$ approach to approximate inverse scattering at xed energy in three dimensions. *Inter- national Math. Research Papers*, 6:287 349, 2005.

[27] J. Nocedal, Updating quasi-Newton matrices with limited storage, *Mathematics of Comp.*, V.35, N.151, 773–782, 1991.

[28] O. Pironneau. *Optimal shape design for elliptic systems*, Springer Verlag, Berlin, 1984.

[29] V. G. Romanov, *Inverse Problems of Mathematical Physics*, VNU, Utrecht, The Netherlands, 1986

[30] A. N. Tikhonov and V. Y. Arsenin. *Solutions of Ill-Posed Problems.* Winston and Sons, Washington, DC, 1977.