SYNTHESIS OF GLOBAL CONVERGENCE AND ADAPTIVITY FOR A HYPERBOLIC COEFFICIENT INVERSE PROBLEM IN 3D *

LARISA BEILINA† AND MICHAEL V. KLIBANOV ‡

Abstract. A globally convergent numerical method for a multidimensional Coefficient Inverse Problem for a hyperbolic equation is presented. Since this method does not use least squares objective functionals, the phenomenon of local minima is avoided. The global convergence is analytically established. It is shown that this technique provides a good first guess for the adaptivity method, which entails to a synthesis of both approaches. Numerical results for the 3-D case are presented.

Key words. two-stage numerical procedure, globally convergent numerical method, adaptive finite element method

AMS subject classifications. 15A15, 15A09, 15A23

1. Introduction. A conventional way to numerically solve a Coefficient Inverse Problem (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, unless a good first guess about this coefficient is known. At the same time, in many important applications a good approximation for the unknown coefficient is unavailable.

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 of the correct solution of that CIP, regardless on the availability of a priori given good approximation 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.

This paper is a continuation of our previous publication [5], where a new globally convergent numerical method for the same CIP was presented. In this method local minima are avoided because a least squares objective functional is not constructed. A fine structure of the PDE operator is used instead. There are five main new ingredients here compared with [5]: (1) We complement our globally convergent method by a locally convergent finite element adaptive technique (below "adaptivity" for brevity), see [7, 8, 9, 10] for this technique. We choose the starting point for the adaptivity the solution obtained by our globally convergent method. (2) The globally convergent

*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.

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

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

algorithm of this publication is different from one in [5] in the sense that, unlike [5], when solving certain nonlinear elliptic equations, we now use "inner" iterations with respect to k for functions $q^k_{n,1}$ until they converge. Consequently, the formulation of the global convergence Theorem 6.1 is different from one of [5]. In particular, the convergence estimate for the unknown coefficient here is obtained under less stringent conditions than one in [5]. Specifically, conditions (6.19) and (6.20) are less restrictive ones than corresponding conditions in [5]. Most importantly, while the convergence estimate in [5] is independent on the iteration number, it is shown here how this estimate actually depends on this number. An advantage of this dependence is that it provides a qualitative explanation of one of observations of numerical studies. All these require a major change in the proof of Theorem 6.1 compared with [5]. (4) The stopping rule differs from one of [5] in the sense that we now evaluate certain L_2 norms at the boundary rather than inside of the domain of interest. (5) Numerical experiments are conducted in the 3-D rather than in the 2-D case of [5].

One of the main ideas of this paper is to synthesize our globally convergent method with a locally convergent adaptivity method. In this synthesis the solution obtained by the former serves as a good first guess for the latter. Indeed, the convergence estimate for the "global part" (Theorem 6.1) depends on a small positive parameter η. This parameter incorporates both the error in the boundary data and the error generated by some approximations of our method. While the error in the boundary data models the error in measurements and is, therefore inevitable, approximation errors of our concrete numerical method are not parts of previously developed locally convergent algorithms. On the other hand, the key to this synthesis is that since η is small, then Theorem 6.1 guarantees that the globally convergent part provides a good approximation for the correct solution. Therefore, it is reasonable to enhance the computed solution via a subsequent application of 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.

In fact, a combination of the globally convergent method of [5] with the adaptivity was recently proposed in the work [6] of the authors. There are two main differences between the current publication and [6]. First, the convergence theorem now is stronger and requires a different proof, since the corresponding theorem of [6] is an extension of the result of [5] on the case of above mentioned inner iterations for functions $q^k_{n,1}$. Second, numerical experiments here are in 3-D, unlike the 2-D case of [6].

The adaptivity technique minimizes least squares objective functionals on a sequence of locally adaptively refined meshes in a series of steps until images are stabilized (usually 3-4 refined meshes are used). The minimization is performed via the quasi-Newton method. Various convergence results for Newton-like methods for general ill-posed problems can be found in the book [3]. The main 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 a certain range). An important point here is that those subdomains are found without a priori knowledge of the solution. It was also shown in [6] that, under certain conditions, the gradient of the Lagrangian can provide an estimate for the accuracy of the reconstruction for the unknown coefficient. 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 [11] that the adaptivity cannot provide good quality images unless a good first guess about the solution

is a priori known. This is because the quasi-Newton method is a locally convergent one. Hence, a synthesis of the adaptivity with the globally convergent method of [5] can 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 [13, 14, 15, 24, 25, 26]. Unlike the current paper, they work for some CIPs with the data resulting from multiple measurements, i.e., either with many positions of the point source or many directions of the initializing plane wave. These works are discussed in [5].

An outline of this work is the following: 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. The idea of this transformation has roots in the method of Carleman estimates for proofs of uniqueness and stability results for CIPs [20, 21]. Since this transformation was described in several previous publications [20, 21], we outline it only briefly here for the sake of completeness. It is the numerical solution of this equation, which represents the major difficulty. The method of this solution was the main new point of [5] and it represents the major departure from the previously developed the so-called "convexification" technique [20]. In section 4 we formulate the layer stripping procedure with respect to s. In section 5 we describe the algorithm. In section 6 a global convergence theorem is proven. In section 7 we describe the computations of the forward problem. In section 8 we briefly outline the adaptivity technique referring to more details to [8, 9, 10]. In section 9 numerical experiments are presented. We summarize our results in section 10.

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 here only because an analogue of the asymptotic behaviour (2.9) is not proved in this case, since (2.9) is actually derived from Theorem 4.1 of [29]. That theorem establishes a certain asymptotic behaviour of the fundamental solution of a hyperbolic equation near the characteristic cone.

Consider the Cauchy problem for the hyperbolic equation

$$c(x)\, u_{tt} = \Delta u \text{ in } \mathbb{R}^3 \times (0, \infty), \tag{2.1}$$

$$u(x, 0) = 0, u_t(x, 0) = \delta(x - x_0). \tag{2.2}$$

Equation (2.1) 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 is $c(x) = \varepsilon_r(x)$, where $\varepsilon_r(x)$ is the spatially distributed dielectric constant of the medium, see [16], 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 (2.1) is such that

$$c(x) \in [d_1, 2d_2], d_1 < d_2, c(x) = 2d_1 \text{ for } x \in \mathbb{R}^3 \setminus \Omega, \tag{2.3}$$

$$c(x) \in C^2\left(\mathbb{R}^3\right), \tag{2.4}$$

We consider the following

Inverse Problem. Suppose that the coefficient $c(x)$ satisfies (2.3) and (2.4), 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 $x_0 \notin \overline{\Omega}$

$$u(x,t) = g(x,t), \forall (x,t) \in \partial\Omega \times (0,\infty). \tag{2.5}$$

The reason why we assume here that the source $x_0 \notin \overline{\Omega}$ is that we do not want to deal with singularities near the source location. 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 (2.1) 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. Practical measurements are calculated 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 $(0,\infty)$.

Uniqueness theorem for this inverse problem is a long standing and well known open question. This question is addressed positively via Carleman estimates only in the case when the $\delta-$ function in (2.2) is replaced with a function which is non vanishing in the entire domain $\overline{\Omega}$ [20, 21]. It is an opinion of the authors that because of applications, it is worthy to study numerical methods for this CIP assuming that uniqueness holds.

Consider the Laplace transform of the functions u,

$$w(x,s) = \int_0^\infty u(x,t)e^{-st}dt, \text{ for } s > \underline{s} = const. > 0, \tag{2.6}$$

where $\underline{s}$ is a certain number. It is sufficient to choose $\underline{s}$ such that the integral (2.6) would converge together with corresponding (x,t)-derivatives. We call the parameter s *pseudo frequency.* The equation for the function w is

$$\Delta w - s^2 c(x) w = -c(x_0)\delta(x - x_0)\, c(x_0), \forall s \geq \underline{s} = const. > 0 \tag{2.7}$$

with the following condition at the infinity

$$\lim_{|x|\to\infty} w(x,s) = 0, \forall s \geq \underline{s} = const. > 0. \tag{2.8}$$

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 behaviour takes place (see Lemma 2.1 in [5])

$$D_x^\beta D_s^\gamma w(x,s) = D_x^\beta D_s^\gamma \left\{ \frac{\exp[-sl(x,x_0)]}{f(x,x_0)} \left[1 + O\left(\frac{1}{s}\right) \right] \right\}, s \to \infty, \tag{2.9}$$

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

$$c(x)U_t = \Delta U - a(x)U, \tag{2.10}$$

$$U(x, 0) = \delta(x - x_0).$$

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 (2.8). In the electromagnetic case equation (2.10) 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].

3. Nonlinear Integral Differential Equation Without the Unknown Coefficient. It follows from (2.6), (2.7) 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 (2.6) and (2.8) lead to

$$\Delta v + |\nabla v|^2 = s^2 c(x) \quad \text{in } \Omega, \tag{3.1}$$

$$v(x, s) = \ln \varphi(x, s), \ \forall (x, s) \in \partial\Omega \times [\underline{s}, \overline{s}], \tag{3.2}$$

where $\varphi(x, s)$ is the Laplace transform (2.6) of the function $g(x, t)$. We eliminate the coefficient $c(x)$ from equation (3.1) via the differentiation with respect to s, since $\partial_s c(x) = 0$. To "isolate" the unknown coefficient in (3.1), introduce a new function

$$H(x, s) = \frac{v}{s^2}. \tag{3.3}$$

It follows from (3.3) and (2.9) that

$$D_x^\alpha(H) = O\left(\frac{1}{s}\right), D_x^\alpha D_s(H) = O\left(\frac{1}{s^2}\right), s \to \infty. \tag{3.4}$$

By (3.1)

$$\Delta H + s^2 (\nabla H)^2 = c(x). \tag{3.5}$$

Denote

$$q(x, s) = \partial_s H(x, s). \tag{3.6}$$

By (3.4) and (3.6)

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

We truncate this integral as

$$H(x,s) \approx -\int_s^{\overline{s}} q(x,\tau)\,d\tau + V(x,\overline{s}), \tag{3.7}$$

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

$$V(x,\overline{s}) \approx H(x,\overline{s}) = \frac{\ln w(x,\overline{s})}{\overline{s}^2}. \tag{3.8}$$

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 (3.4) 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 (3.5)- (3.7) the following (approximate) integral nonlinear differential equation

$$\Delta q - 2s^2 \nabla q \cdot \int_s^{\overline{s}} \nabla q(x,\tau)\,d\tau + 2s\left[\int_s^{\overline{s}} \nabla q(x,\tau)\,d\tau\right]^2 \tag{3.9}$$

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

In addition, (3.2), (3.3) and (3.6) imply that the following Dirichlet boundary condition is given for the function q

$$q(x,s) = \psi(x,s),\ \forall (x,s) \in \partial\Omega \times [\underline{s},\overline{s}], \tag{3.10}$$

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 $D_x^\alpha q, |\alpha| \leq 2$. Then the corresponding approximation for the target coefficient can be found via (3.5) as

$$c(x) = \Delta H + \underline{s}^2 (\nabla H)^2, \tag{3.11}$$

where the function H is approximated via (3.7). Although any value of the pseudo frequency $s \in [\underline{s},\overline{s}]$ can be used in (3.11), 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 (3.9) 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 (3.9), we find updates for V using solutions of forward problems (2.1), (2.2) and the formula (3.8). In those forward problems corresponding approximations for the unknown coefficient c, obtained from (3.11), are used. We refer to subsection 5.4 of [19] for the first procedure of working with tails, which was applied to a linearized CIP.

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,

$$\int_s^{\overline{s}} \nabla q(x,\tau) d\tau = (s_{n-1} - s)\nabla q_n(x) + h\sum_{j=1}^{n-1} \nabla q_j(x), s \in (s_n, s_{n-1}). \tag{4.1}$$

We approximate the boundary condition (3.10) as a piecewise constant function,

$$q_n(x) = \overline{\psi}_n(x), x \in \partial\Omega, \tag{4.2}$$

where

$$\overline{\psi}_n(x) = \frac{1}{h}\int_{s_n}^{s_{n-1}} \psi(x,s)\, ds. \tag{4.3}$$

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

$$C_{n,\lambda}(s) = e^{\lambda(s - s_{n-1})}, \quad s \in (s_n, s_{n-1}], \tag{4.4}$$

and integrate with respect to s over (s_n, s_{n-1}). In (4.4) $\lambda >> 1$ 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])

$$\begin{aligned} L_n(q_n) &:= \Delta q_n - A_{1,n}\left(h\sum_{i=1}^{n-1}\nabla q_i\right)\nabla q_n + A_{1n}\nabla q_n \nabla V - \varepsilon q_n \\ &= 2\frac{I_{1,n}}{I_0}(\nabla q_n)^2 - A_{2,n}h^2\left(\sum_{i=1}^{n-1}\nabla q_i(x)\right)^2 \\ &+ 2A_{2,n}\nabla V\left(h\sum_{i=1}^{n-1}\nabla q_i\right) - A_{2,n}(\nabla V)^2, n = 1, ..., N, \end{aligned} \tag{4.5}$$

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 (4.2), (4.5) for a nonlinear elliptic PDE with the unknown function $q_n(x)$. In (4.5) the tail function V is also unknown. An important observation is that

$$\frac{|I_{1,n}(\lambda, h)|}{I_0(\lambda, h)} \leq \frac{4\overline{s}^2}{\lambda}. \tag{4.6}$$

Therefore, by taking $\lambda >> 1$, we mitigate the influence of the nonlinear term with $(\nabla q_n)^2$ in (4.5). This enables us to solve each elliptic Dirichlet boundary value problem (4.2), (4.5) iteratively at each n via solving a linear problem on each step. Theorem 6.1 assures convergence of this procedure. We have added the term $-\varepsilon q_n$ to the left hand side of equation (4.5), 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

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

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. Below $C^{k+\alpha}(\overline{\Omega})$ are Hőlder spaces, where $k \geq 0$ is an integer and $\alpha \in (0,1)$ [22]. Denote $|f|_{k+\alpha} = \|f\|_{C^{k+\alpha}(\overline{\Omega})}, \forall f \in C^{k+\alpha}(\overline{\Omega})$. Our algorithm reconstructs iterative approximations $c_{n,i}(x) \in C^{\alpha}(\overline{\Omega})$ 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 (2.1), (2.2). 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 $\widehat{c}_{n,k} \in C^{\alpha}(\mathbb{R}^3)$, $\widehat{c}_{n,k} \geq d_1$ 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 convergences of certain "sub-procedures". Specifications of corresponding convergence criteria are given in subsection 9.1.

In accordance with (3.7), (3.11) and (4.1) denote

$$H_{n,i}(x) = hq_{n,i} + h\sum_{j=1}^{n-1} q_j(x) + V_{n,i}(x), x \in \Omega, \tag{5.1}$$

$$c_{n,i}(x) = \Delta H_{n,i} + s_n^2 (\nabla H_{n,i})^2, \tag{5.2}$$

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 $i = 1, ..., m_n$. We set

$$q_0 := 0, q_{1,1}^0 := 0, V_{1,1}(x) := V_{1,1}^0(x), \tag{5.3}$$

$$q_{n,1}^0 := q_{n-1}, V_{n,1}(x) := V_{n-1,m_{n-1}}(x), \text{for } n \geq 2, \tag{5.4}$$

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 (5.3), (5.4) . We now construct the function $q_{n,1}$. To do this, we solve iteratively the following Dirichlet boundary value problems

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

$$2\frac{I_{1n}}{I_0}\left(\nabla q_{n,1}^{k-1}\right)^2 - A_{2n}h^2\left(\sum_{j=1}^{n-1}\nabla q_j(x)\right)^2 \tag{5.5}$$

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

$$q_{n,1}^k(x) = \overline{\psi}_n(x), x \in \partial\Omega. \tag{5.6}$$

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 (5.1), (5.2) 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 (2.1), (2.2) with $c(x) := \widehat{c}_{n,1}(x)$. We obtain the function $u_{n,1}(x,t)$. Calculate the Laplace transform (2.6) of this function and obtain the function $w_{n,1}(x, \overline{s})$ this way. Next, following (3.8), we set for $x \in \Omega$

$$V_{n,2}(x, \overline{s}) = \frac{1}{\overline{s}^2}\ln w_{n,1}(x, \overline{s}) \in C^{2+\alpha}\left(\overline{\Omega}\right). \tag{5.7}$$

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

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

$$= 2\frac{I_{1n}}{I_0}\left(\nabla q_{n,i-1}\right)^2 - A_{2n}h^2\left(\sum_{j=1}^{n-1}\nabla q_j(x)\right)^2 \tag{5.8}$$

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

$$q_{n,i}(x) = \overline{\psi}_n(x), x \in \partial\Omega. \tag{5.9}$$

Having the function $q_{n,i}$, we reconstruct the next approximation $c_{n,i} \in C^\alpha(\overline{\Omega})$ for the target coefficient using (5.1), (5.2), 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 (2.1), (2.2) with $c(x) := \widehat{c}_{n,i}(x)$, calculate the Laplace transform (2.6) and update the tail as in (5.7), 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

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

$$V_{n+1,1}(x,\overline{s}) = \frac{1}{\overline{s}^2} \ln w_{n,m_n}(x,\overline{s}) \in C^{2+\alpha}\left(\overline{\Omega}\right). \tag{5.10}$$

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 (5.7) that in principle, to update the tail, one can solve the problem (2.7), (2.8) for $s = \overline{s}$ instead of the problem (2.1), (2.2). However, our computational experience shows that it is better to proceed via solving the problem (2.1), (2.2) 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 (2.4), and this function is an exact solution of our Inverse Problem with the "ideal" exact data in $g^*(x,t)$ in (2.5). The Laplace transform (2.6) 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 $w^*(x,s) \in C^{2+\alpha}\left(\mathbb{R}^3 \setminus \{|x - x_0| < \gamma\}\right), \forall \gamma > 0, \forall s \geq \underline{s}$ is the solution of the forward problem (2.7), (2.8) 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 (3.9) with the following boundary condition (see (3.10)

$$q^*(x,s) = \psi^*(x,s), (x,s) \in \partial\Omega \times [\underline{s}, \overline{s}], \tag{6.1}$$

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 (3.9), (3.10) with the *exact boundary* condition $\psi^*(x,s)$.

Hence,

$$q^*(x,s) \in C^{2+\alpha}\left(\overline{\Omega}\right) \times C^{\infty}\left[\underline{s},\overline{s}\right]. \tag{6.2}$$

We now follow (4.1)-(4.5). First, we approximate functions $q^*(x,s)$ and $\psi^*(x,s)$ in (6.1) and (6.2) via piecewise constant functions with respect to $s \in [\underline{s},\overline{s}]$. Let

$$q_n^*(x) = \frac{1}{h}\int\limits_{s_n}^{s_{n-1}} q^*(x,s)\,ds,\ \overline{\psi}_n^*(x) = \frac{1}{h}\int\limits_{s_n}^{s_{n-1}} \psi^*(x,s)\,ds. \tag{6.3}$$

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

$$|Q_n(x,s)|_{2+\alpha} \leq C^* h, |\Psi_n(x,s)|_{2+\alpha} \leq C^* h, n = 1,...,N, \text{ for } s \in [s_n, s_{n-1}], \tag{6.4}$$

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 $C^{2+\alpha}\left(\overline{\Omega}\right) \times C^1[\underline{s},\overline{s}]$ norm of the function $q^*(x,s)$. Hence, we can assume that

$$\max_{1\leq n\leq N} |q_n^*|_{2+\alpha} \leq C^*. \tag{6.5}$$

Without a loss of generality, we assume that

$$C^* \geq 1. \tag{6.6}$$

By the Tikhonov concept, the constant C^* should be known a priori. By (2.9), it is reasonable to assume that C^* is independent on $\overline{s}$, although we do not use this assumption. By (6.3)

$$q_n^*(x) = \overline{\psi}_n^*(x), x \in \partial\Omega. \tag{6.7}$$

Hence we obtain the following analogue of equation (4.5)

$$\Delta q_n^* - A_{1,n}\left(h\sum_{i=1}^{n-1}\nabla q_i^*(x)\right)\nabla q_n^* + A_{1,n}\nabla q_n^*\nabla V^*$$

$$= 2\frac{I_{1,n}}{I_0}\left(\nabla q_n^*\right)^2 - A_{2,n}h^2\left(\sum_{i=1}^{n-1}\nabla q_i^*(x)\right)^2 \tag{6.8}$$

$$+2A_{2,n}\nabla V^*\left(h\sum_{i=1}^{n-1}\nabla q_i^*(x)\right) - A_{2,n}\left|\nabla V^*\right|^2 + F_n(x,h,\lambda),$$

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

$$\max_{\lambda h\geq 1} |F_n(x,h,\lambda)|_{\alpha} \leq C^* h. \tag{6.9}$$

We also assume that the function $g(x,t)$ in (2.8) is given with an error. This naturally produces an error in the function $\psi(x,s)$ in (3.10). An additional error is introduced due to the averaging in (4.3). Hence, it is reasonable to assume that

$$\left\| \overline{\psi}_n^*(x) - \overline{\psi}_n(x) \right\|_{C^{2+\alpha}(\partial\Omega)} \leq C^*(\sigma + h), \tag{6.10}$$

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. Global 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

$$\overline{s} > 1, \ \lambda h \geq 1, \tag{6.11}$$

it was shown in [5] that

$$\max_{1 \leq n \leq N} \{|A_{1,n}| + |A_{2,n}|\} \leq 8\overline{s}^2. \tag{6.12}$$

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

$$M^* = 2C^* \max\left(8\overline{s}^2, \max_{1 \leq n \leq N} \{|A_{1,n}| + |A_{2,n}|\}\right). \tag{6.13}$$

Hence, (6.12) and (6.13) imply that

$$M^* = 16C^*\overline{s}^2. \tag{6.14}$$

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

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

By the Schauder theorem, there exists unique solution $u \in C^{2+\alpha}(\overline{\Omega})$ of this boundary value problem, and with a constant $K = K(\Omega) > 1$, depending only on the domain Ω, the following estimate holds

$$|u|_{2+\alpha} \leq K\left[\|g\|_{C^{2+\alpha}(\partial\Omega)} + |f|_\alpha\right]. \tag{6.15}$$

In the formulation of Theorem 6.1 we provide estimates (6.19)-(6.24) via M^* and also use (6.14) to obtain estimates via $\overline{s}$. Note that the definition of the norm in the space $C^\alpha(\overline{\Omega})$ implies that

$$|f_1 f_2|_\alpha \leq |f_1|_\alpha |f_2|_\alpha, \ \forall f_1, f_2 \in C^\alpha\left(\overline{\Omega}\right). \tag{6.16}$$

Theorem 6.1. *Let $\Omega \subset \mathbb{R}^3$ be a convex bounded domain with the boundary $\partial\Omega \in C^3$. Suppose that inequalities (6.6) and (6.11) hold. Let the exact coefficient $c^*(x) \in C^2(\mathbb{R}^3), c^* \in [2d_1, 2d_2]$ and $c^*(x) = 2d_1$ 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 $w_c(x, \overline{s}) \in C^3\left(\mathbb{R}^3 \setminus \{|x - x_0| < \gamma\}\right), \forall \gamma > 0$ of the problem (2.6), (2.7). Let $V_c(x) = \overline{s}^{-2} \ln w_c(x, \overline{s}) \in C^{2+\alpha}\left(\overline{\Omega}\right)$ be the corresponding tail function. Suppose that the cut-off pseudo frequency $\overline{s}$ is so large that for any such function $c(x)$ the following estimates hold*

$$|V^*|_{2+\alpha} \leq \xi, |V_c|_{2+\alpha} \leq \xi, \tag{6.17}$$

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

$$|V_{1,1}|_{2+\alpha} \leq \xi. \tag{6.18}$$

Denote $\eta := 2(h + \sigma + \xi + \varepsilon)$. Let K be the constant of the Schauder theorem and $\overline{N} \leq N$ be the total number of functions q_n calculated by the algorithm of section 5. 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

$$\beta \leq \frac{1}{384 K \overline{s}^2} = \frac{1}{24 K M^*}. \tag{6.19}$$

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

$$\eta \leq \eta_0(K, C^*, d_1, \overline{s}) = \min\left(\frac{1}{16 K M^*}, \frac{3}{8} d_1\right) = \min\left(\frac{1}{256 K C^* \overline{s}^2}, \frac{3}{8} d_1\right), \tag{6.20}$$

$$\lambda \geq \lambda_0(C^*, K, \overline{s}, \eta) = \max\left(\frac{(C^*)^2}{4}, 48 K C^* \overline{s}^2, \frac{1}{\eta^2}\right). \tag{6.21}$$

Then for each appropriate n the sequence $\{q_{n,1}^k\}_{k=1}^\infty$ converges in $C^{2+\alpha}\left(\overline{\Omega}\right)$ and the following estimates hold

$$|q_n - q_n^*|_{2+\alpha} \leq 2 K M^* \left(\frac{1}{\sqrt{\lambda}} + 3\eta\right), n \in \left[1, \overline{N}\right], \tag{6.22}$$

$$|q_n|_{2+\alpha} \leq 2C^*, \ n \in \left[1, \overline{N}\right], \tag{6.23}$$

$$|c_n - c^*|_\alpha \leq \frac{\eta}{2 \cdot 9^{n-1}} + \frac{23}{8}\eta, \ n \in \left[2, \overline{N}\right]. \tag{6.24}$$

In addition, functions $c_{n,k}(x), \widehat{c}_{n,k}(x) \geq d_1$ *in* Ω *and* $\widehat{c}_{n,k}(x) = 2d_1$ *outside of* Ω.

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. For example, 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 (2.1) and the whole space $\mathbb{R}^3$ only for the sake of a convenient formulation of the asymptotic behaviour (2.9). 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. In the spirit of the previous remark, in actual computations the estimate (6.24) is a qualitative one in the sense that it is more likely that in the reality

$$|c_n - c^*|_\alpha \leq \frac{C}{A^{n-1}}\eta + B\eta, \ n \in [2, \overline{N}]$$

with some positive numbers B, C and with a number $A > 1$. Even though it is shown in the course of the proof of this theorem that $|c_n - c^*|_\alpha \leq 8/3 \cdot \eta$, which is stronger, the estimate (6.24) is still quite helpful for an explanation of numerical results, see subsection 9.3. Indeed, (6.24) shows that on first iterations the error in our solution decreases rapidly (the first term in the right hand side of (6.24). On the other hand, this error cannot be made less than a constant (23/8 in our case) times the small parameter η. This parameter incorporates both the error in our model and the error in the boundary data.

3. 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.

4. 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 [17]. Therefore, we have a vector $(\overline{s}, \overline{N}, \lambda, \varepsilon_n)$ of regularization parameters. Setting $\overline{N}(h)\, h = \beta = const. > 0$ is in an agreement with, e.g., Lemma 6.2 on page 156 of [17], 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 number β is small because our algorithm is originated by equation (3.9), 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.

We assume below in this section that conditions of Theorem 6.1 hold. We obtain from (6.6), (6.20) and (6.21) that

$$\frac{C^*}{2\sqrt{\lambda}} \leq 1, \frac{1}{\sqrt{\lambda}} + 3\eta \leq \frac{C^*}{2KM^*}, \frac{KM^*}{\lambda} \leq \frac{1}{3}, \frac{1}{\sqrt{\lambda}} \leq \eta, \tag{6.25}$$

Denote

$$\widetilde{q}^k_{n,1} = q^k_{n,1} - q^*_n, \widetilde{q}_{n,i} = q_{n,i} - q^*_n,$$

$$\widetilde{V}_{n,k} = V_{n,k} - V^*, \widetilde{c}_{n,k} = c_{n,k} - c^*, \widetilde{\psi}_n = \overline{\psi}_n - \overline{\psi}^*_n$$

$$\widetilde{H}_{n,i}(x) = H_{n,i}(x) - H^*(x, s_n), \widetilde{H}_n(x) = H_n(x) - H^*(x, s_n),$$

where $H^*(x,s)$ is the function $H_n(x)$ in (5.1) in the case when functions q_j and V_n are replaced with q_j^* and V^* respectively.

6.3. Proof of Theorem 6.1. This proof basically consists in estimating norms $\left|\widetilde{q}_{n,1}^k\right|_{2+\alpha}, |\widetilde{q}_{n,i}|_{2+\alpha}$ from the above. First, we estimate norms $\left|\widetilde{q}_{1,k}^1\right|_{2+\alpha}$. By (6.17) and (6.18)

$$\left|\widetilde{V}_{1,1}\right|_{2+\alpha} \leq 2\xi \leq \eta. \tag{6.26}$$

Substituting $n = 1$ in (6.8), subtracting it from (5.5) and subtracting (6.7) from (5.6), we obtain

$$\Delta\widetilde{q}_{1,1}^k - \varepsilon\widetilde{q}_{1,1}^k + A_{1,1}\nabla V_{1,1}\nabla\widetilde{q}_{1,1}^k = -2\frac{I_{1,1}}{I_0}\nabla\widetilde{q}_{1,1}^{k-1}\left(\nabla q_{1,1}^{k-1} + \nabla q_1^*\right) \tag{6.27}$$

$$-A_{1,1}\nabla\widetilde{V}_{1,1}\nabla q_1^* - A_{2,1}\nabla\widetilde{V}_{1,1}\left(\nabla V_{1,1} + \nabla V^*\right) + \varepsilon q_1^* - F_1,$$

$$\widetilde{q}_{1,1}^1(x) = \widetilde{\psi}_1(x), x \in \partial\Omega. \tag{6.28}$$

By (6.20) $\varepsilon \leq \eta/2 < 1$. Also, since $K, C^* > 1$, then by (6.12), (6.18) and (6.20)

$$|A_{1,1}\nabla V_{1,1}| \leq 4\overline{s}^2\eta \leq \frac{1}{64K} < 1. \tag{6.29}$$

Hence, combining the Schauder theorem (6.15) with (4.6), (6.9), (6.11)-(6.14), (6.17), (6.20) and (6.26)-(6.29), we obtain

$$\left|\widetilde{q}_{1,1}^k\right|_{2+\alpha} \leq \frac{KM^*}{2C^*\lambda}\left|\widetilde{q}_{1,1}^{k-1}\right|_{1+\alpha}\left|q_{1,1}^{k-1} + q_1^*\right|_{1+\alpha} + 3KM^*\eta. \tag{6.30}$$

First, let $k = 1$. Since by (5.3) $q_{1,1}^0 = 0$, then $\widetilde{q}_{1,1}^0 = -q_1^*$. By (6.5) $|\nabla q_1^*|_\alpha^2 \leq (C^*)^2$. Hence, (6.30) implies that

$$\left|\widetilde{q}_{1,1}^1\right|_{2+\alpha} \leq KM^*\left[\frac{C^*}{2\lambda} + 3\eta\right].$$

Hence, using the first inequality (6.25), we obtain

$$\left|\widetilde{q}_{1,1}^1\right|_{2+\alpha} \leq KM^*\left(\frac{1}{\sqrt{\lambda}} + 3\eta\right) \leq 2KM^*\left(\frac{1}{\sqrt{\lambda}} + 3\eta\right).$$

Hence, the second inequality (6.25) and (6.5) imply that

$$\left|q_{1,1}^1\right|_{2+\alpha} \leq \left|\widetilde{q}_{1,1}^1\right|_{2+\alpha} + |q^*|_{2+\alpha} \leq 2C^*. \tag{6.31}$$

Assume now that

$$\left|\widetilde{q}_{1,1}^{k-1}\right|_{2+\alpha} \leq 2KM^* \left(\frac{1}{\sqrt{\lambda}} + 3\eta\right), k \geq 2. \tag{6.32}$$

Then similarly with (6.31)

$$\left|q_{1,1}^{k-1}\right|_{2+\alpha} \leq 2C^*. \tag{6.33}$$

We obtain from (6.27), (6.28), (6.32) and (6.33)

$$\left|\widetilde{q}_{1,1}^{k}\right|_{2+\alpha} \leq \frac{3\left(KM^*\right)^2}{\lambda}\left(\frac{1}{\sqrt{\lambda}} + 3\eta\right) + 3KM^*\eta.$$

Hence, the third inequality (6.25) leads to

$$\left|\widetilde{q}_{1,1}^{k}\right|_{2+\alpha} \leq 2KM^* \left(\frac{1}{\sqrt{\lambda}} + 3\eta\right), k \geq 1. \tag{6.34}$$

Hence, we obtain similarly with (6.31) that

$$\left|q_{1,1}^{k}\right|_{2+\alpha} \leq 2C^*, k \geq 1. \tag{6.35}$$

Estimates (6.34) and (6.35) enable us to prove convergence of functions $q_{1,1}^k$ for $k \to \infty$. Let $m, r > 2$ be two positive integers. Denote $a_{m,r} = q_{1,1}^m - q_{1,1}^r$. Then $a_{m,r} = \widetilde{q}_{1,1}^m - \widetilde{q}_{1,1}^r$. Set in (6.27) $k := m$ and then $k := r$. Next, subtract two resulting equations, and use the following

$$\nabla\widetilde{q}_{1,1}^{m-1}\left(\nabla q_{1,1}^{m-1} + \nabla q_1^*\right) - \nabla\widetilde{q}_{1,1}^{r-1}\left(\nabla q_{1,1}^{r-1} + \nabla q_1^*\right)$$

$$= \nabla\widetilde{q}_{1,1}^{m-1}\left(\nabla q_{1,1}^{m-1} + \nabla q_1^*\right) - \nabla\widetilde{q}_{1,1}^{r-1}\left(\nabla q_{1,1}^{m-1} + \nabla q_1^*\right)$$

$$+\nabla\widetilde{q}_{1,1}^{r-1}\left(\nabla q_{1,1}^{m-1} + \nabla q_1^*\right) - \nabla\widetilde{q}_{1,1}^{r-1}\left(\nabla q_{1,1}^{r-1} + \nabla q_1^*\right)$$

$$= \nabla a_{m-1,r-1}\left(\nabla q_{1,1}^{m-1} + \nabla q_1^*\right) + \nabla\widetilde{q}_{1,1}^{r-1} \cdot \nabla a_{m-1,r-1}$$

$$= \nabla a_{m-1,r-1}\left(\nabla q_{1,1}^{m-1} + \nabla\widetilde{q}_{1,1}^{r-1} + \nabla q_1^*\right).$$

We obtain

$$\Delta a_{m,r} - \varepsilon a_{m,r} + A_{1,1}\nabla V_{1,1}\nabla a_{mr} =$$

$$-2\frac{I_{1,1}}{I_0}\nabla a_{m-1,r-1} \cdot \left(\nabla q_{1,1}^{m-1} + \nabla\widetilde{q}_{1,1}^{r-1} + \nabla q_1^*\right),$$

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

Hence, by the Schauder theorem (6.15), second and third inequalities (6.25), (6.34), (6.35) and (4.6)

$$|a_{m,r}|_{2+\alpha} \leq \frac{KM^*}{\lambda} |a_{m-1,r-1}|_{2+\alpha} \leq \frac{1}{3} |a_{m-1,r-1}|_{2+\alpha}. \tag{6.36}$$

It follows from (6.36) that the sequence $\{q_{1,1}^k\}_{k=1}^{\infty}$ satisfies the Cauchy convergence criterion. Convergence of other sequences $\{q_{n,1}^k\}_{k=1}^{\infty}$ can be proven similarly. Thus, these proofs are omitted below.

Since functions $\widetilde{q}_{1,1}$ and $q_{1,1}$ are estimated via (6.34) and (6.35), we now can estimate the norm $|\widetilde{c}_{1,1}|_{\alpha}$. To do this, we note that it follows from (5.1) and (5.2) that

$$|\widetilde{c}_{1,1}|_{\alpha} \leq \left|\widetilde{H}_{1,1}\right|_{2+\alpha} \left[1 + \overline{s}^2 \left(|H_{1,1}|_{2+\alpha} + |H_1^*|_{2+\alpha}\right)\right].$$

By (5.1), (6.17)-(6.19) and the fourth inequality (6.25)

$$\left|\widetilde{H}_{1,1}\right|_{2+\alpha} \leq 8KM^*\beta\eta + \eta \leq \frac{1}{3}\eta + \eta = \frac{4}{3}\eta.$$

Next, (5.1), (6.18), (6.19), (6.20) and (6.35) lead to

$$1 + \overline{s}^2 \left(|H_{1,1}|_{2+\alpha} + |H_1^*|_{2+\alpha}\right) \leq 1 + 3\overline{s}^2 C^*\beta + \overline{s}^2\eta \leq 1 + \frac{M^*}{5}\beta + \frac{1}{256}$$

$$\leq 1 + \frac{1}{120} + \frac{1}{256} < 2.$$

Thus, the last two inequalities and (6.20) imply that

$$|\widetilde{c}_{1,1}|_{\alpha} \leq \frac{8}{3}\eta \leq d_1. \tag{6.37}$$

Since $c^* \geq 2d_1$, then (6.37) implies that $c_{1,1}(x) \geq d_1$ in Ω. Hence, the function $\widehat{c}_{1,1} \in C^{\alpha}\left(\mathbb{R}^3\right), \widehat{c}_{1,1}(x) \geq d_1$ in Ω and $\widehat{c}_{1,1}(x) = 2d_1$ in $\mathbb{R}^3 \setminus \Omega$ (see beginning of section 5). This, along with one of conditions of Theorem 6.1, ensures that $|V_{1,2}|_{2+\alpha} \leq \xi$. Hence, one can prove similarly with the above that estimates (6.34), (6.35) and (6.37) are valid for functions $\widetilde{q}_{1,2}, q_{1,2}$ and $\widetilde{c}_{1,2}$ respectively. To do this, one should use (5.8) and (5.9) at $n = 1, i = 2$. Repeating this process m_1 times, we obtain the same estimates for functions $\widetilde{q}_1, q_1, \widetilde{c}_1$. In addition, we also obtain that $c_1(x) \geq d_1$ in Ω, $\widehat{c}_1 \in C^{\alpha}\left(\mathbb{R}^3\right), \widehat{c}_1(x) \geq d_1$ in Ω and $\widehat{c}_1(x) = 2d_1$ in $\mathbb{R}^3 \setminus \Omega$. Hence, $|V_{2,1}|_{2+\alpha} \leq \xi$.

Assume now that

$$|\widetilde{q}_j|_{2+\alpha} \leq 2KM^* \left(\frac{1}{\sqrt{\lambda}} + 3\eta\right), j = 1, ..., n-1, \tag{6.38}$$

$$|q_j|_{2+\alpha} \leq 2C^*, j = 1, ..., n-1, \tag{6.39}$$

$$c_j(x) \geq d_1 \text{ in } \Omega,\ \widehat{c}_j(x) \geq d_1 \text{ in } \mathbb{R}^3,\ j = 1, ..., n-1. \tag{6.40}$$

We now obtain these estimates at $j = n$. It follows from (6.40) and (6.17) that

$$|V_{n,1}|_{2+\alpha} \leq \xi \leq \frac{\eta}{2}, \quad \left|\widetilde{V}_{n,1}\right| \leq 2\xi \leq \eta. \tag{6.41}$$

For brevity consider only functions $q_{n,i}$ with $i \geq 2$, since the case of $q^k_{n,1}$ can be considered very similarly, given convergence of the sequence $\left\{q^k_{n,1}\right\}$. Thus, we assume that estimates (6.38) and (6.39) hold for the function $q_{n,1}$.

Subtracting (6.8) from (5.8) and (6.7) from (5.9), we obtain for $i \geq 2$

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

$$= 2\frac{I_{1,n}}{I_0}\left[\nabla\widetilde{q}_{n,i-1}\left(\nabla q_{n,i-1} + \nabla q^*_n\right)\right]$$

$$+\left(A_{1,n}\nabla q^*_n - A_{2,n}h\sum_{j=1}^{n-1}\left(\nabla q_j + \nabla q^*_j\right) + 2A_{2,n}\nabla V_{n,i}\right)\left(h\sum_{j=1}^{n-1}\nabla\widetilde{q}_j\right) \tag{6.42}$$

$$+\left[2A_{2,n}h\sum_{j=1}^{n-1}\nabla q^*_j - A_{1,n}\nabla q^*_n - A_{2,n}\left(\nabla V_{n,i} + \nabla V^*\right)\right]\nabla\widetilde{V}_{n,i} + \varepsilon q^*_n - F_n,$$

$$\widetilde{q}_{n,i}\,|_{\partial\Omega} = \widetilde{\psi}_n(x). \tag{6.43}$$

We begin estimate the sum of 2^{nd}, 3^{rd}, 4^{th} and 5^{th} terms in the right hand side of (6.42). As to the second term, using (6.5), (6.12)-(6.14), (6.20), (6.39) and (6.41), we obtain

$$\left|A_{1,n}\nabla q^*_n - A_{2,n}h\sum_{j=1}^{n-1}\left(\nabla q_j + \nabla q^*_j\right) + 2A_{2,n}\nabla V_{n,1}\right|_\alpha \leq \frac{M^*}{2} + \frac{3M^*\beta}{2} + \frac{M^*}{4} = M^*\left(1 + \frac{3}{2}\beta\right).$$

On the other hand, by (6.38)

$$h\sum_{j=1}^{n-1}|\nabla\widetilde{q}_j|_\alpha \leq 2KM^*\beta\left(\frac{1}{\sqrt{\lambda}} + 3\eta\right). \tag{6.44}$$

Hence, (6.16) implies that

$$\left|A_{1,n}\nabla q^*_n - A_{2,n}h\sum_{j=1}^{n-1}\left(\nabla q_j + \nabla q^*_j\right) + 2A_{2,n}\nabla V_{n,1}\right|_\alpha \cdot \left|h\sum_{j=1}^{n-1}\nabla\widetilde{q}_j(x)\right|_\alpha \tag{6.45}$$

$$\leq 2K\left(M^*\right)^2\beta\left(1+\frac{3}{2}\beta\right)\left(\frac{1}{\sqrt{\lambda}}+3\eta\right).$$

Estimate now the sum of 3^{rd}, 4^{th} and 5^{th} terms in the right hand side of (6.42). We obtain similarly with the above

$$\left|\left(2A_{2,n}h\sum_{j=1}^{n-1}\nabla q_j^*-A_{1,n}\nabla q_n^*-A_{2,n}\left(\nabla V_{n,1}+\nabla V^*\right)\right)\nabla\widetilde{V}_{n,1}+\varepsilon q_n^*-F_1\right|_\alpha\leq 2M^*\left(1+\frac{\beta}{2}\right)\eta. \tag{6.46}$$

Combining this with (6.44), (6.45) and (6.16), we obtain the following estimate

$$\left|\left(A_{1,n}\nabla q_n^*-A_{2,n}h\sum_{j=1}^{n-1}\left(\nabla q_j+\nabla q_j^*\right)+2A_{2,n}\nabla V_{n,1}\right)\left(h\sum_{j=1}^{n-1}\nabla\widetilde{q}_j\right)\right|_\alpha$$

$$+\left|\left(2A_{2,n}h\sum_{j=1}^{n-1}\nabla q_j^*-A_{1,n}\nabla q_n^*-A_{2,n}\left(\nabla V_{n,1}+\nabla V^*\right)\right)\nabla\widetilde{V}_{n,1}+\varepsilon q_n^*-F_1\right|_\alpha \tag{6.47}$$

$$\leq 2K\left(M^*\right)^2\beta\left(1+\frac{3}{2}\beta\right)\left(\frac{1}{\sqrt{\lambda}}+3\eta\right)+2M^*\left(1+\frac{\beta}{2}\right)\eta.$$

Since $K, M^*>1$, then (6.19) and the 4^{th} inequality (6.25) imply that

$$2K\left(M^*\right)^2\beta\left(1+\frac{3}{2}\beta\right)\left(\frac{1}{\sqrt{\lambda}}+3\eta\right)\leq 8K\left(M^*\right)^2\beta\left(1+\frac{3}{2}\beta\right)\eta\leq\frac{1}{2}M^*\eta.$$

By (6.19)

$$2M^*\left(1+\frac{\beta}{2}\right)\eta\leq\frac{5}{2}M^*\eta. \tag{6.48}$$

Hence, we obtain from (6.47) and (6.48) that

$$\left|\left(A_{1,n}\nabla q_n^*-A_{2,n}h\sum_{j=1}^{n-1}\left(\nabla q_j+\nabla q_j^*\right)+2A_{2,n}\nabla V_{n,1}\right)\left(h\sum_{j=1}^{n-1}\nabla\widetilde{q}_j\right)\right|_\alpha$$

$$+\left|\left(2A_{2,n}h\sum_{j=1}^{n-1}\nabla q_j^*-A_{1,n}\nabla q_n^*-A_{2,n}\left(\nabla V_{n,1}+\nabla V^*\right)\right)\nabla\widetilde{V}_{n,1}+\varepsilon q_n^*-F_1\right|_\alpha \tag{6.49}$$

$$\leq\frac{1}{2}M^*\eta+\frac{5}{2}M^*\eta=3M^*\eta.$$

It follows from (6.12), (6.14), (6.19), (6.20) and (6.39) that $C^{\alpha}\left(\overline{\Omega}\right)$ norms of coefficients at $\nabla\widetilde{q}_{n,i}, \widetilde{q}_{n,i}$ in equation (6.42) do not exceed 1. Hence, applying the estimate (6.15) to the Dirichlet boundary value problem (6.42), (6.43) and using (4.6), we obtain

$$|\widetilde{q}_{n,i}|_{2+\alpha} \leq \frac{KM^*}{2C^*\lambda} |\nabla\widetilde{q}_{n,i-1}|_{\alpha} |\nabla q_{n,i-1} + \nabla q_n^*|_{\alpha} + 3KM^*\eta.$$

Consider first the case $i = 2$. Since estimates (6.38) and (6.39) hold true for $\widetilde{q}_{n,1}, q_{n,1}$,then (6.5), (6.38), (6.39) and the third inequality (6.25) imply that

$$|\widetilde{q}_{n,2}|_{2+\alpha} \leq \frac{3KM^*}{\lambda} KM^* \left(\frac{1}{\sqrt{\lambda}} + 3\eta\right) + 3KM^*\eta \leq 2KM^* \left(\frac{1}{\sqrt{\lambda}} + 3\eta\right),$$

which establishes (6.38) for the function $\widetilde{q}_{n,2}$. Hence, similarly with (6.31) we obtain $|q_{n,2}|_{2+\alpha} \leq 2C^*$, which proves (6.39) for $q_{n,2}$. Using (6.38) and (6.39), we obtain similarly with (6.37) that $|\widetilde{c}_{n,i}|_{\alpha} \leq 8/3 \cdot \eta \leq d_1$. Hence, functions $\widehat{c}_{n,2}(x), c_{n,2}(x) \geq d_1$ in Ω and $\widehat{c}_{n,2}(x) = 2d$ in $\mathbb{R}^3 \setminus \Omega$, which establishes (6.40) for functions $c_{n,2}, \widehat{c}_{n,2}(x)$. The latter, (5.10) and one of conditions of this theorem guarantee that $|V_{n,3}|_{2+\alpha} \leq \xi$. Recalling that $q_n = q_{n,m_n}$ and applying the mathematical induction principle, we obtain that estimates (6.38)-(6.40) are valid for $j = n$.

Having estimates (6.38)-(6.40) for $j = 1, ..., n$, we now obtain the estimate (6.24). Denote

$$p_n := \sum_{j=1}^{n} |\widetilde{q}_j|_{2+\alpha}, g_n = hp_n, n \in \left[2, \overline{N}\right].$$

It follows from the above proof that

$$\left| \left(A_{1,n}\nabla q_n^* - A_{2,n}h\sum_{j=1}^{n-1} \left(\nabla q_j + \nabla q_j^*\right) + 2A_{2,n}\nabla V_{n,1} \right) \left(h\sum_{j=1}^{n-1} \nabla\widetilde{q}_j \right) \right|_{\alpha} \leq M^* \left(1 + \frac{3}{2}\beta\right) hp_{n-1}$$

$$\leq 2M^*hp_{n-1}.$$

Hence, it follows from (6.46) and (6.48) that the sum of all terms in equation (6.42), excluding the first one can be estimated from the above via $2M^*hp_{n-1} + 5/2 \cdot M^*\eta$. First, consider the case when in (6.42) $\widetilde{q}_{n,i}$ is replaced with $\widetilde{q}_{n,1}^k$ and respectively $\widetilde{q}_{n,i-1}$ is replaced with $\widetilde{q}_{n,1}^{k-1}$. Since the sequence $\left\{q_{n,1}^k\right\}_{k=1}^{\infty}$ converges, we can replace in (6.42) the vector $\left(\widetilde{q}_{n,k}^k, \widetilde{q}_{n,1}^k, q_{n,1}^{k-1}\right)$ with the vector $(\widetilde{q}_{n,1}, \widetilde{q}_{n,1}, q_{n,1})$. Hence, applying to the boundary value problem (6.42), (6.43) the Schauder theorem (6.15), (6.20), (6.21) and the 4^{th} inequality (6.25), we obtain

$$|\widetilde{q}_{n,1}|_{2+\alpha} \leq \frac{|\widetilde{q}_{n,1}|_{2+\alpha}}{4} + 2KM^*hp_{n-1} + \frac{5}{2}KM^*\eta.$$

Or

$$|\widetilde{q}_{n,1}|_{2+\alpha} \leq \frac{8}{3}KM^*hp_{n-1} + \frac{10}{3}KM^*\eta.$$

Similarly we obtain the same estimate for $|\widetilde{q}_{n,i}|_{2+\alpha}$. Hence, recalling that $\widetilde{q}_n = \widetilde{q}_{n,m_n}$, we obtain

$$|\widetilde{q}_n|_{2+\alpha} \leq \frac{8}{3}KM^*hp_{n-1} + \frac{10}{3}KM^*\eta. \tag{6.50}$$

Substituting in (6.50) $\widetilde{q}_{n-k}$ for $\widetilde{q}_n$, we obtain the following sequence of estimates

$$|\widetilde{q}_{n-k}|_{2+\alpha} \leq \frac{8}{3}KM^*hp_{n-k-1} + \frac{10}{3}KM^*\eta,\ 0 \leq k \leq n-2. \tag{6.51}$$

Summing up all estimates (6.51) for functions $\widetilde{q}_{n-k}$ with $0 \leq k \leq n-2$, we obtain

$$p_n - |\widetilde{q}_1|_{2+\alpha} \leq \frac{8}{3}KM^*h\sum_{i=1}^{n-1} p_i + \frac{10}{3}KM^*n\eta.$$

Since $|p_i|_{2+\alpha} \leq |p_{i+1}|_{2+\alpha}$ and $h\overline{N} = \beta$, then

$$|p_n|_{2+\alpha} \leq \frac{8}{3}KM^*\beta p_{n-1} + \frac{10}{3}KM^*\overline{N}\eta + |\widetilde{q}_1|_{2+\alpha}.$$

Hence, multiplying by h and using (6.38) and the 4^{th} inequality (6.25), we obtain

$$g_n \leq \frac{8}{3}KM^*\beta g_{n-1} + \frac{10}{3}KM^*\beta\eta + 4KM^*\eta^2.$$

Hence, (6.19) and (6.20) imply that

$$g_n \leq \frac{1}{9}g_{n-1} + \frac{7}{18}\eta, n \in \left[2, \overline{N}\right].$$

Iterating this inequality and using the formula for the sum of the geometrical progression, we obtain

$$g_n \leq \frac{1}{9^{n-1}}g_1 + \frac{7}{16}\eta, n \in \left[2, \overline{N}\right].$$

Since $g_1 = h\,|\widetilde{q}_1|_{2+\alpha} \leq |\widetilde{q}_1|_{2+\alpha}\,\eta/2$, then (6.20), (6.38) and the 4^{th} inequality (6.25) imply that

$$g_n \leq \frac{\eta}{4 \cdot 9^{n-1}} + \frac{7}{16}\eta, n \in \left[2, \overline{N}\right]. \tag{6.52}$$

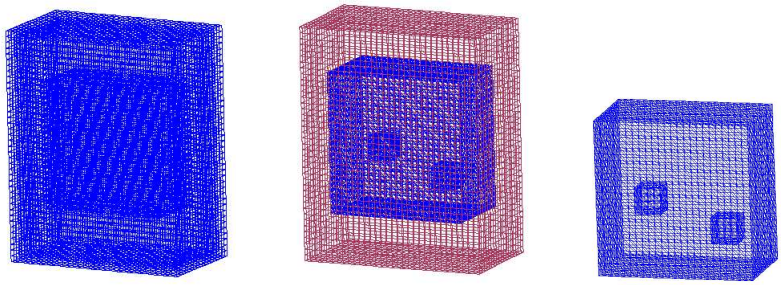
We now derive the estimate (6.24). Repeating the above arguments, which were given for $|\widetilde{c}_{1,1}|_\alpha$, we obtain

$$|\widetilde{c}_n|_\alpha \leq \left|\widetilde{H}_n\right|_{2+\alpha}\left[1 + \overline{s}^2\left(|H_n|_{2+\alpha} + |H_n^*|_{2+\alpha}\right)\right] \leq 2\left|\widetilde{H}_n\right|_{2+\alpha}. \tag{6.53}$$

Also, by (5.1) and (6.17) $\left|\widetilde{H}_n\right|_{2+\alpha} \leq g_n + \eta$. Hence, it follows from (6.53) that $|\widetilde{c}_n|_\alpha \leq 2\,(g_n + \eta)$. Combining this with (6.52), we obtain

$$|\widetilde{c}_n|_\alpha \leq \frac{1}{2 \cdot 9^{n-1}}\eta + \frac{23}{8}\eta, n \in \left[2, \overline{N}\right].$$

□



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

FIG. 7.1. *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 hexahedron* Ω *and* $c(x) = 1$ *for* $x \in G \setminus \Omega$.

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 (7.2) with the given function $c(x)$. To solve the forward problem, we use the hybrid FEM/FDM method described in [12]. The computational domain in all our tests $G = G_{FEM} \cup G_{FDM}$ is set as $G = [-4.0, 4.0] \times [-5.0, 5.0] \times [-2.5, 2.0]$. This domain is split into a finite element domain $G_{FEM} := \Omega = [-3.0, 3.0] \times [-3.0, 3.0] \times [-2.0, 1.5]$ and a surrounding domain G_{FDM} with a structured mesh, see Figure 7.1. The space mesh in Ω consists of tetrahedrons and in G_{FDM} - of cubes, with the mesh size $\tilde{h} = 0.25$ in the overlapping regions. At the top and bottom boundaries of G we use first-order absorbing boundary conditions [18] 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 $G \subset \mathbb{R}^3$ (Figure 7.1). The coefficient $c(x)$ is unknown only in domain $\Omega \subset G$ and

$$c(x) = 1 \text{ in } G \setminus \Omega. \tag{7.1}$$

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 7.1 and ∂G_3 is the union of left,right, front and back sides of this domain. In our tests the forward problem is

$$\begin{aligned}
c(x)\frac{\partial^2 u}{\partial t^2} - \triangle u &= 0, \quad \text{in } G \times (0, T), \\
u(\cdot, 0) = 0, \ \frac{\partial u}{\partial t}(\cdot, 0) &= 0, \ \text{in } G, \\
\partial_n u\big|_{\partial G_1} &= f(t), \ \text{on } \partial G_1 \times (0, t_1], \\
\partial_n u\big|_{\partial G_1} &= \partial_t u, \ \text{on } \partial G_1 \times (t_1, T), \\
\partial_n u\big|_{\partial G_2} &= \partial_t u, \ \text{on } \partial G_2 \times (0, T), \\
\partial_n u\big|_{\partial G_3} &= 0, \ \text{on } \partial G_3 \times (0, T),
\end{aligned} \tag{7.2}$$

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

$$f(t) = \frac{(\sin(\overline{s}t - \pi/2) + 1)}{10},\ 0 \le t \le t_1 := \frac{2\pi}{\overline{s}}, T = 8.22t_1.$$

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 [18] 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 our computations the upper limit of the integral in the Laplace transform (2.6) is T.

8. The Adaptivity Technique. In this section we briefly describe the adaptivity technique referring for details to [8, 9, 10]. To use the adaptivity technique, we formulate the inverse problem for the boundary value problem (7.2) as an optimization problem, where we seek the unknown coefficient $c(x)$, which gives the solution of the boundary value problem (7.2) for the function $u(x,t)$ with the best least squares fit to the time domain observations $g(x,t)$, see (2.5). 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

$$E(u,c) = \frac{1}{2}\int\limits_{S_T} (u\mid_{S_T} - g(x,t))^2 d\sigma dt + \frac{1}{2}\gamma \int\limits_{\Omega} (c - c_0)^2 \ dx, \tag{8.1}$$

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 refinements are used, the function $c_0(x)$ is again taken from the globally convergent stage. 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 $G \setminus \Omega$, then given the function $g(x,t) = u\mid_{\partial\Omega}$, one can uniquely determine the function $u(x,t)$ for $(x,t) \in (G \setminus \Omega) \times (0,T)$ as the solution of the boundary value problem for equation (7.2) with boundary conditions on both boundaries ∂G and $\partial\Omega$. Hence, one can uniquely determine the function $p(x,t)$,

$$\frac{\partial u}{\partial n}\mid_{S_T} = p(x,t). \tag{8.2}$$

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 (7.2) with the correct value of the coefficient $c(x)$.

Denote

$$\begin{aligned}
H_u^2(Q_T) &= \{f \in H^2(Q_T) : f(x,0) = f_t(x,0) = 0\},\\
H_u^1(Q_T) &= \{f \in H^1(Q_T) : f(x,0) = 0\},\\
H_\varphi^2(Q_T) &= \{f \in H^2(Q_T) : f(x,T) = f_t(x,T) = 0\},\\
H_\varphi^1(Q_T) &= \{f \in H^1(Q_T) : f(x,T) = 0\},\\
U &= H_u^2(Q_T) \times H_\varphi^2(Q_T) \times C^2(\overline{\Omega}),\\
\bar{U} &= H_u^1(Q_T) \times H_\varphi^1(Q_T) \times L_2(\Omega),\\
\bar{U}^1 &= L_2(Q_T) \times L_2(Q_T) \times L_2(\Omega),
\end{aligned} \tag{8.3}$$

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 $\bar{U}^1$.

To solve the problem of the minimization of the functional (8.1), we introduce the Lagrangian

$$L(v) = E(u,c) + \int\limits_{Q_T} \varphi \cdot (cu_{tt} - \Delta u)\, dxdt, \forall \varphi \in H^2_\varphi(Q_T), \tag{8.4}$$

where $\varphi \in H^2_\varphi(Q_T)$ is the Lagrange multiplier and $v = (u, \varphi, c) \in U$. Since the function u solves equation (7.2), then $L(v) = E(u,c)$. This is because the second term in $L(v)$ is zero. Integration by parts and (8.4) leads to

$$L(v) = E(u,c) - \int\limits_{Q_T} c(x) u_t \varphi_t dxdt + \int\limits_{Q_T} \nabla u \nabla \varphi dxdt - \int\limits_{S_T} p\varphi d\sigma dt. \tag{8.5}$$

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

$$L'(v)(\overline{v}) = 0, \quad \forall \bar{v} = (\bar{u}, \bar{\varphi}, \bar{c}) \in \bar{U} \tag{8.6}$$

where $L'(v)(\cdot)$ is the Frechet derivative of L at the point v. We obtain that (8.6) is equivalent with (see details in [6])

$$cu_{tt} - \triangle u = 0, (x,t) \in Q_T, \tag{8.7}$$

$$u(x,0) = u_t(x,0) = 0, \tag{8.8}$$

$$\partial_n u \mid_{S_T} = p(x,t); \tag{8.9}$$

$$c\varphi_{tt} - \triangle \varphi = 0, (x,t) \in Q_T, \tag{8.10}$$

$$\varphi(x,T) = \varphi_t(x,T) = 0, \tag{8.11}$$

$$\frac{\partial \varphi}{\partial n} \mid_{S_T} = (u-g)(x,t), (x,t) \in S_T; \tag{8.12}$$

$$\gamma(c - c_0) - \int_0^T u_t \varphi_t \, dt = 0, x \in \Omega. \tag{8.13}$$

The boundary value problem (8.10)-(8.12) is the adjoint problem to (8.7)-(8.9) and should be solved backwards in time. Uniqueness and existence theorems for initial boundary value problems (8.7)-(8.9) and (8.10)-(8.12), including the case of weak $H^1_u(Q_T)$ and $H^1_\varphi(Q_T)$ 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 (8.7)-(8.9) and (8.10)-(8.12). We find weak solutions of these problems via the FEM.

To formulate the FEM for boundary value problems (8.7)-(8.9) and (8.10)-(8.12), we introduce finite element spaces $W^u_h \subset H^1_u(Q_T)$ and $W^\varphi_h \subset H^1_\varphi(Q_T)$ 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 $U_h = W_h^p \times W_h^\varphi \times V_h \subset \bar{U}$. The space U_h as a discrete analogue of the space $\bar{U}$. In order to use below one of results of [6], which provides an error bound for the unknown coefficient, we set the norm in U_h to be same as the norm in $\bar{U}^1$. However, if restricting attention to a more conventional case of estimating the error in the Lagrangian only, one can set the norm in U_h to be the same as the norm in $\bar{U}$. The functional $L(v_h)$ is defined in terms of (8.5) and $L'(v)(\overline{v})$ is defined in terms of (8.6). The FEM for (8.6) now reads: Find $v_h \in U_h$, such that

$$L'(v_h)(\overline{v}) = 0, \quad \forall \overline{v} \in U_h. \tag{8.14}$$

We solve this discrete problem using the quasi-Newton method with the limited storage [27], see details for our specific implementation in [6].

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 (8.6) and v_h is a solution of the problem (8.14). 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 neighbourhood of the exact solution $v^* \in U$ of our original CIP. Thus, we assume that

$$\|v - v^*\|_{\bar{U}} \leq \delta \text{ and } \|v - v_h\|_{\bar{U}} \leq \delta, \tag{8.15}$$

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

$$\begin{aligned} L(v) - L(v_h) &= \int_0^1 \frac{d}{d\epsilon} L(\epsilon v + (1-\epsilon)v_h) d\epsilon \\ &= \int_0^1 L'(\epsilon v + (1-\epsilon)v_h)(v - v_h)\, d\epsilon = L'(v_h)(v - v_h) + R, \end{aligned} \tag{8.16}$$

where R is the second order, with respect to $v - v_h$, remainder term, $|R| \leq C\|v - v_h\|_{\bar{U}}^2$ with a certain positive constant C. We ignore R because of (8.15).

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 (8.14) with the splitting $v - v_h = (v - v_h^I) + (v_h^I - v_h)$, we obtain the following error representation:

$$L(v) - L(v_h) \approx L'(v_h)(v - v_h^I), \tag{8.17}$$

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 of the adaptivity idea, because it allows one to use the Galerkin orthogonality (8.14) 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 [8]-[10]. It turns out that the dominating contribution of the error in the Lagrangian (8.4) 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 the unknown coefficient $c(x)$, i.e. the left hand side of (8.13), attains its maximum. The latter forms the basis for the adaptivity technique.

8.2. The adaptive algorithm. In this section we present our adaptive algorithm based on computations of the residuals for the computed coefficient $c(x)$. See more details for a posteriori error estimate for the unknown coefficient in [10, 6].

Let M be the dimension of the finite dimensional space U_h and $\{\psi_k\}_{k=1}^M$ be an orthonormal basis in U_h. Assuming that a solution of a certain dual problem for the Hessian of the Lagrangian exists, it was shown in [6] that the following approximate aposteriori error estimate for the computed coefficient is valid

$$\left\| c_h^I - c_h \right\|_{L_2(\Omega)} \leq MC_2 A(\Omega) \max_K |[\widetilde{c}_h]| \int_0^T \left(\max_{\overline{\Omega}} R_{c_1}(x,t) + \max_{\overline{\Omega}} R_{c_2}(x,t) \right) dt, \tag{8.18}$$

where $[\widetilde{c}_h]$ on a space element K denotes the maximum of the modulus of the jump of the $\widetilde{c}_h$ across a face of K, the maximum is taken over all space elements, $R_{c_1}(x,t) = \left| \frac{\partial \varphi_h}{\partial t} \right| \cdot \left| \frac{\partial u_h}{\partial t} \right|$, $R_{c_2} = \gamma |(c_h - c_0)|$, $A(\Omega)$ is the area of the domain Ω (volume in the 3-d case) and C_2 is an interpolation constant. See more details for arguments in [6]. If, however, solution of the the dual problem for the Hessian of Lagrangian does not exist for some (or all) functions ψ_k, then it follows from (8.17) that the integral term in (8.18) estimates from the above the error in the Lagrangian,

$$|L(v) - L(v_h)| \approx |L'(v_h)\left(v - v_h^I\right)| \leq C_2 A(\Omega) \max_K |[c_h]| \int_0^T \left(\max_{\overline{\Omega}} R_{c_1}(x,t) + \max_{\overline{\Omega}} R_{c_2}(x,t) \right) dt. \tag{8.19}$$

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 (8.18) and (8.19) allow us to control the error either in the computed reconstructed coefficient c_h or in the Lagrangian. In our computations we use the following version of the adaptive algorithm.

Adaptive algorithm

0. Choose an initial mesh K_h and an initial time partition J_0 of the time interval $(0,T]$. Start with the 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 (8.7)-(8.9) on K_h and J_k, with $c(x) = c^m$.
2. Compute the solution φ^m of the adjoint problem (8.10)-(8.12) 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, see details in [6, 28]
$$c^{m+1} = c^m - \alpha H^m g^m.$$
4. Stop computing c if either the norm of the gradient g^m of the Lagrangian with respect to the coefficient in (8.13) is $||g^m||_{L_2(\Omega)} < \theta$ or norms $||g^n||_{L_2(\Omega)}$ are stabilized or start to grow. 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 residuals, R_{c_1}, R_{c_2} and refine the mesh at all points where
$$\int_0^T \left(\max_{\overline{\Omega}} R_{c_1}(x,t) + \max_{\overline{\Omega}} R_{c_2}(x,t) \right) dt > tol. \tag{8.20}$$
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 an initial guess $c_0 = c_{glob}$ 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 cubes, where $c(x) = 4$, see Figure 7.1-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 cubes, although, following the Tikhonov concept (as mentioned in section 2), we have assumed the knowledge of the constant $d_1 = 1/2$, see (2.3) and (7.1). Because of this, the starting value for the tail $V_{1,1}(x, \overline{s})$ was computed via solving the forward problem (7.2) 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 (3.8), we took $V_{1,1}(x, \overline{s}) = \overline{s}^{-2} \ln w_{c\equiv 1}(x, \overline{s})$.

We have found that the pseudo frequency interval $s \in [3.3, 4.3]$ was the optimal one for the above domains G, Ω (section 7). The step size in the s-direction was chosen as $h = 0.05$. Hence, $N = 20$ in our case. We have chosen the same sequence ε_n of regularization parameters as in [5]. However, unlike [5], where the sequence λ_n was chosen also, we choose here the parameter λ independent on n, see below.

Once the function q_n is calculated, we update the function $c := c_n$ as in (5.1), (5.2), 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
$$\text{either } F_n^k \geq F_n^{k-1} \text{ or } F_n^k \leq \varepsilon,$$

it. n $\lambda = 50$	i=1	i=2	i=3	it. n $\lambda = 200$	i=1	i=2	i=3
1	0.0522995	0.0522995		1	0.052307	0.052307	
2	0.0523043	0.0521799		2	0.0523043	0.0521758	
3	0.0535235	0.053353		3	0.0535235	0.053353	
4	0.0516891	0.0556757		4	0.0516891	0.0556757	
5	0.0467661	0.091598		5	0.0467661	0.091598	
6	0.0466467	0.0440336	0.0464053	6	0.0466467	0.0440336	0.0464053
7	0.048653	0.0658041		7	0.048651	0.0658031	
8	0.0631709	0.0893371		8	0.0631753	0.0893179	
9	0.0851995	0.112022		9	0.085511	0.112321	
10	0.0914011	0.106414		10	0.0915352	0.10644	
11	0.0900873	0.104467		11	0.0905234	0.104808	
12	0.111039	0.133793		12	0.111136	0.134055	
13	0.141459	0.167344		13	0.141494	0.166125	
14	0.176421	0.219103		14	0.174968	0.222117	
15	0.238352	0.296523		15	0.240944	0.29716	
16	0.327406	0.463613		16	0.328997	0.464465	
17	0.528386	0.606531		17	0.53069	0.606824	
18	0.630857	0.680105		18	0.630438	0.681458	
19	0.704224	0.747767		19	0.705058	0.748638	

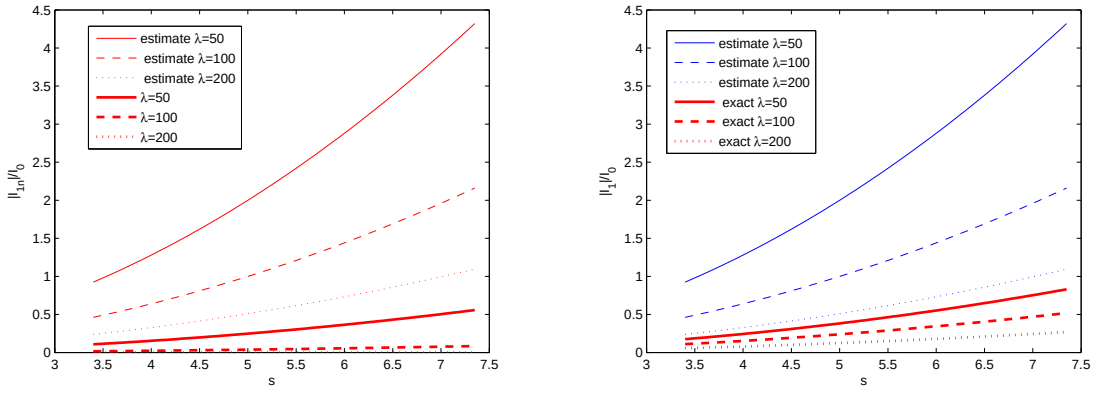
TABLE 9.1

Test 1. Computed L_2-norms of the $F_{n,i} = ||q_{n,i}|_{\partial\Omega} - \overline{\psi}_n||_{L_2(\partial\Omega)}$. with $\lambda = 50$ and $\lambda = 200$.

it. n	i=1	i=2	i=3
1	0.0522995	0.0522995	
2	0.0523043	0.0521772	
3	0.0535235	0.053353	
4	0.0516891	0.0556757	
5	0.0467661	0.091598	
6	0.0466467	0.0440336	0.0464053
7	0.0486575	0.0657632	
8	0.0631762	0.0892608	
9	0.0852419	0.111969	
10	0.0914603	0.106285	
11	0.090428	0.104433	
12	0.11104	0.133783	
13	0.141538	0.166841	
14	0.175796	0.221395	
15	0.239744	0.299591	
16	0.329464	0.465232	
17	0.530307	0.606822	
18	0.63031	0.681715	
19	0.704982	0.747981	

TABLE 9.2

Test 1. Computed L_2 norms of the $F_{n,i} = ||q_{n,i}|_{\partial\Omega} - \overline{\psi}_n||_{L_2(\partial\Omega)}$ with $\lambda = 100$.



(a) (b)

FIG. 9.1. *Test 1. On (a) we show computed and on (b) exact values of* $|I_{1,n}|/I_0$ *compared with the estimation* $|I_{1,n}|/I_0 \leq 4\bar{s}^2/\lambda$ *(see (4.6)) for different values of* λ. *Computations performed on pseudo-frequency interval* $[3.4, 7.4]$ *with* $h = 0.05$. *Here, x-axis presents pseudo-frequency interval. One can see that (4.6) significantly overestimates the value of* $|I_{1,n}|/I_0$.

where $\varepsilon = 0.001$ is a small tolerance number of our choice. In other words, we stop iterations, when either F_n^k start 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

$$F_{n,i} \geq F_{n,i-1} \tag{9.1}$$

or

$$F_{n,i} \leq \varepsilon, \tag{9.2}$$

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 (9.1)-(9.2) 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 procedure allows us to get a more flexible stopping rule in the globally convergent algorithm than one in [5]. 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).

9.2. Test 1. We test the globally convergent method on the reconstruction of the structure given on Figure 7.1-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 $t \in (t_1, T)$.

One of the main steps in the numerical analysis of the problems (4.2)-(4.5) is computations of integrals $I_0, I_{1,n}, A_{1,n}, A_{2,n}$ in (4.5). These integrals can be approximated by either approximate or exact formula. We observe that to compute these integrals, we need to calculate

$$\int_{s_n}^{s_{n-1}} s^k C_{n,\lambda}(s) ds,$$

where k is the degree of s. We have derived the following recurrent formula for numerical calculation of these integrals

$$\int_{s_n}^{s_{n-1}} s^k C_{n,\lambda}(s) ds = \frac{s_{n-1}^k - s_n^k e^{-\lambda h}}{\lambda} - \frac{k}{\lambda} \int_{s_n}^{s_{n-1}} s^{k-1} C_{n,\lambda}(s) ds.$$

Let us define

$$M_k := \int_{s_n}^{s_{n-1}} s^k C_{n,\lambda}(s) ds.$$

Then integrals $I_{1,n}, A_{1,n}, A_{2,n}$ can be computed as

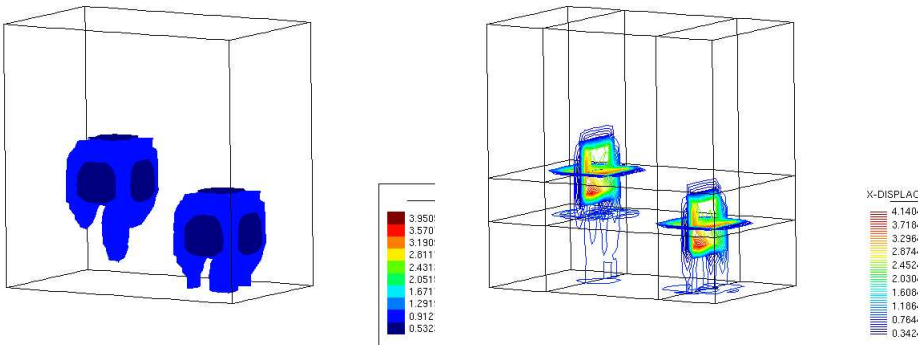
$$\begin{aligned} I_{1,n} &= 3 s_{n-1} M_2 - s_{n-1}^2 M_1 - 2 M_3, \\ A_{1,n} &= \frac{2}{I_0}(3 M_2 - 2 s_{n-1} M_1), \\ A_{2,n} &= \frac{2}{I_0} M_1. \end{aligned}$$

In Tables 9.1, 9.2 we analyze computed L_2 norms of the $F_{n,i}$ for different values of λ in CWF. We observe that significant changes in λ cause only insignificant changes in L_2 norms of the $F_{n,i}$. The results in Tables 9.1, 9.2 are in an agreement with results in Figure 9.1, where in Figure 9.1a we present approximated and on Figure 9.1b exact values of the ratio $|I_{1,n}|/I_0$ compared with the estimate (4.5). Figure 9.1 shows that a significant grow of the value of λ has a very small influence to the value of $|I_{1,n}|/I_0$ on the pseudo-frequency interval $[3.3, 4.3]$ which we take in actual computations. This fact indicates that L_2 norms of the $F_{n,i}$ will be almost unchanged for steeply growing λ. Therefore we can work only with one value of λ for all n.

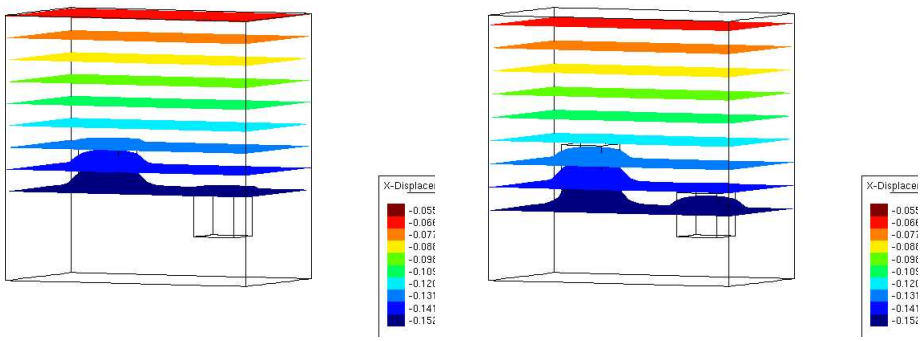
Figures 9.2a,b display the reconstructed coefficient for the "ideal" case when the tail is known exactly. Figures 9.2c-h present isosurfaces of the exact tail for different values of s.

Figures 9.3-9.4 display isosurfaces of resulting images of functions $c_{n,k}, n = 1, ..., 15$ with numerically approximated integrals $I_0, I_{1,n}, A_{1,n}, A_{2,n}$ by midpoint rule, which corresponds to the computed integrals of the Figure 9.1a with $\lambda = 200$. Figures 9.5-9.6 display isosurfaces of resulting images of functions $c_{n,k}, n = 1, ..., 15$ with exact computed integrals $I_0, I_{1,n}, A_{1,n}, A_{2,n}$, which corresponds to the exact integrals of the Figure 9.1b with $\lambda = 200$. 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.

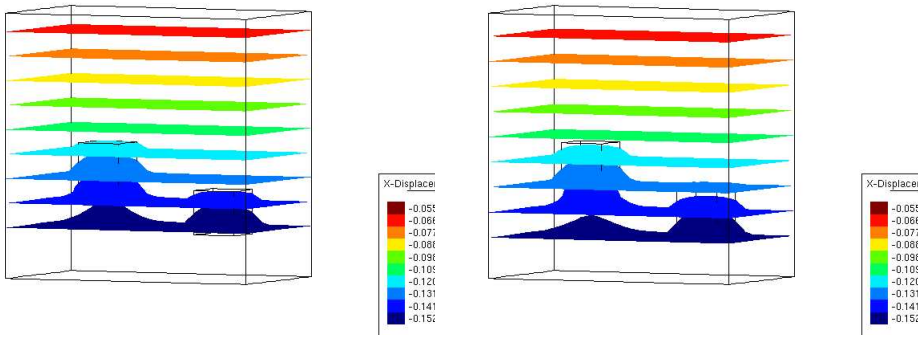
One can see from Tables 9.1, 9.2 that the number m_n of iterations with respect to tails indeed varies with n, since m_n is chosen automatically now, using the criterion (9.1)-(9.2). We observe



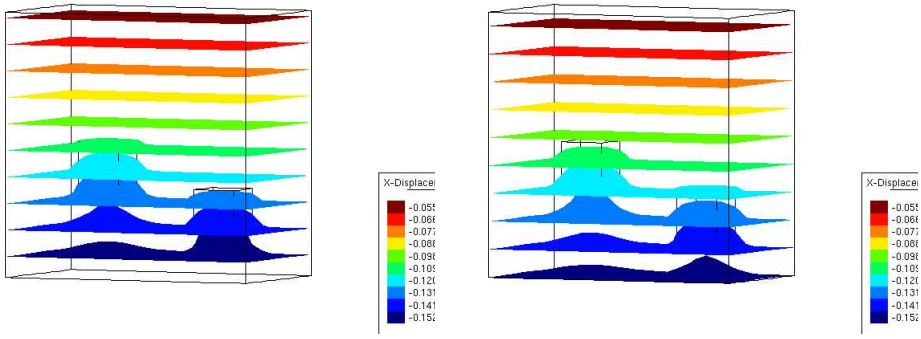
a) c(x) with exact tail b) c(x) with exact tail



c) exact tail s=7.2 d) exact tail s=7.4



e) exact tail s=7.5 f) exact tail s=7.6



g) exact tail s=7.8 h) exact tail s=8.0

FIG. 9.2. *Test 1: on (a)-(b) we show reconstructed coefficient with exact tail; on (c)-(h) we present isosurfaces of the exact tail for different values of s.*

that the computed $F_{n,i}$ decrease until computing the function q_7. Next, $F_{7,2} > F_{6,2}$ and norms $F_{n,i}$ start to grow with the increase of n. They are stabilized for $n = 10, 11$ and then grow steeply for $n = 12, ..., 19$. Thus, we conclude, that $\overline{N} = 7$ and we take $c_{7,1}$ as our final reconstruction result of the first stage of our two-stage procedure.

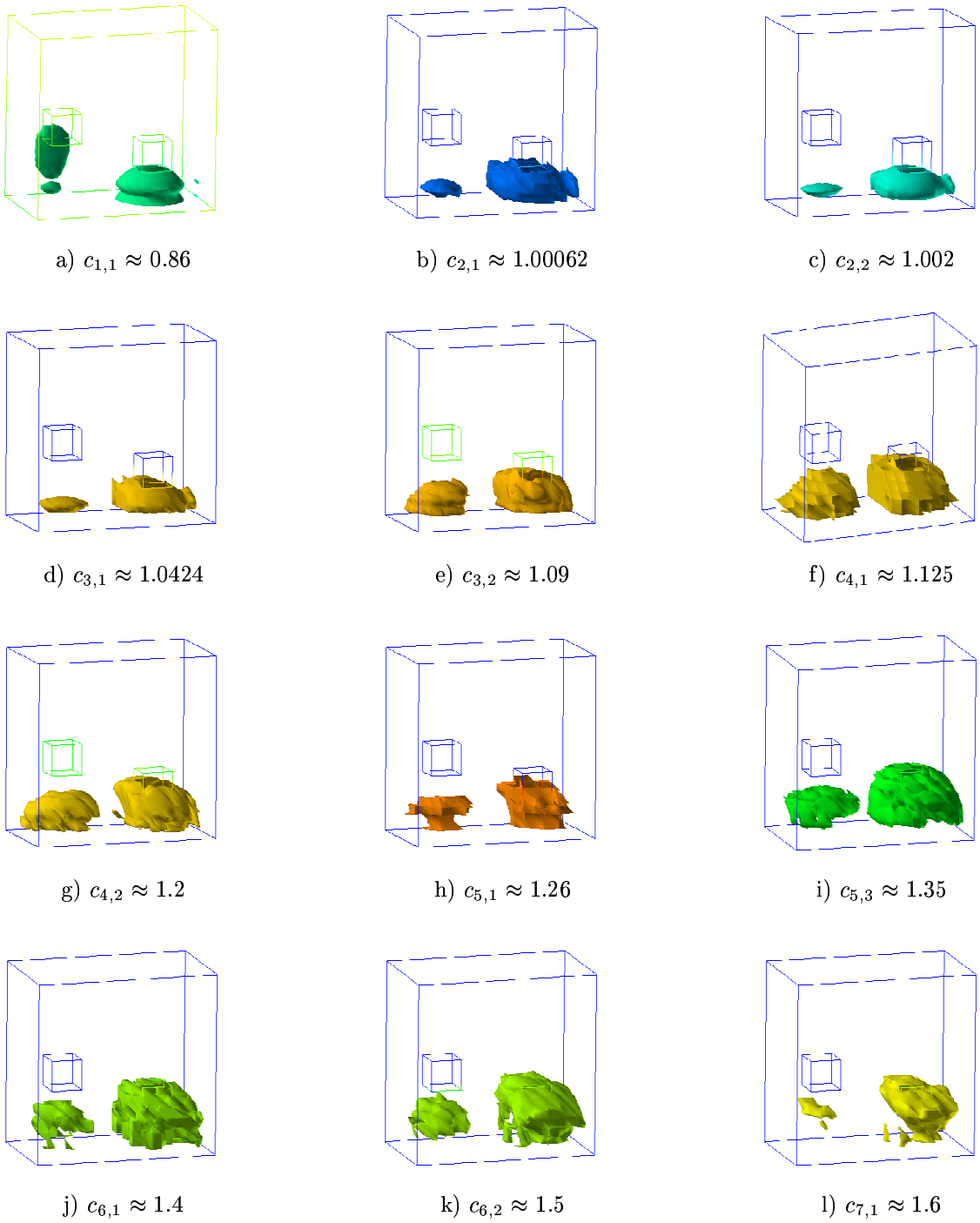
9.3. Correspondence between Tables 9.1, 9.2 and the theory. The behavior of norms $F_{n,i}$ in Tables 9.1 and 9.2 can be qualitatively explained by the estimate (6.24) of Theorem 6.1. Indeed, the variable term in the right hand side of (6.24) says that one should expect that these norms would decrease first with the grow of n for rather small values of n. In particular, we note that these norms do not change when n changes from 1 to 2, which is in a good agreement with (6.24), since $n \geq 2$ in this estimate. Next, after a certain value of n one essentially enters the "level of error" of $\eta \cdot 23/8$, also see the second remark after Theorem 6.1. It is natural therefore that as soon as this level is achieved, one cannot guarantee that further iterations would still lead to a better approximation of the correct solution. This is why norms $F_{n,i}$ start to grow with n for $n \geq 8$. Hence, in our case the level of error is reached at $n = 7$.

The rapid growth of norms $F_{n,i}$ with the increase of n for $n \geq 12$ can also be explained by the fact that the critical level $\beta = \overline{N}h$ in (6.19) of the length of the s-interval is achieved at $n = 12$. Indeed, Theorem 6.1 does not guarantee anything for $n > \overline{N}$. Thus, in our case $\overline{N} = 12, \beta = 12 \cdot 0.05 = 0.6$. Finally, we again refer to Lemma 6.2 on page 156 of [17]. Indeed, it can be derived from this lemma that in the case of ill-posed problems one cannot have too many iterations. The above in turn paves the way for a subsequent application of the adaptivity technique, which is free from such components of η as ε, ξ and h, although is not free from σ.

9.4. The synthesis of the globally convergent algorithm with the adaptivity. The goal of three tests of this subsection is to demonstrate the performance of the synthesis of our globally convergence algorithm with the adaptivity technique. The difference between these three tests is in the noise level which we introduce in the boundary data $g = u\,|_{\partial\Omega}$ for the adaptivity. These are $0\%, 3\%$ and 5% levels of noise. Note that on the globally convergent stage we have introduced only 5% noise in the data. Nevertheless, we use the solution obtained on that stage as the starting point c_0 for the adaptivity for all three levels of noise. In addition, we take the same starting point, although properly interpolated, for all refined meshes. It seems to be at the first glance that it would be better to take on each follow up step the function $c_0(x)$ obtained on the previous step. However, our computational experience has shown that images are worse in this case. In our opinion this points towards the robustness of the globally convergent method. In all tests let Γ be the bottom side of the cube Ω, which is 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. We take the regularization parameter $\gamma = 0.01$ in all our tests.

The adaptive algorithm means that we find the solution of our problem in an iterative process, where we start with a coarse mesh and find an approximate solution by the quasi-Newton method on this mesh, see subsection 8.2. Next, we evaluate residuals as in (8.20). 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 computations again on this new mesh. We stop the iterative process when L_2-norms of the computed gradient for the coefficient either are stabilized or start to grow for all further refinements of the mesh. Let $|R_c(x)| = |R_{c_1}(x)| + |R_{c_2}(x)|$, see (8.19). We refine the mesh in all regions where

$$|R_c(x)| \geq \zeta \max |R_c(x)|, \tag{9.3}$$



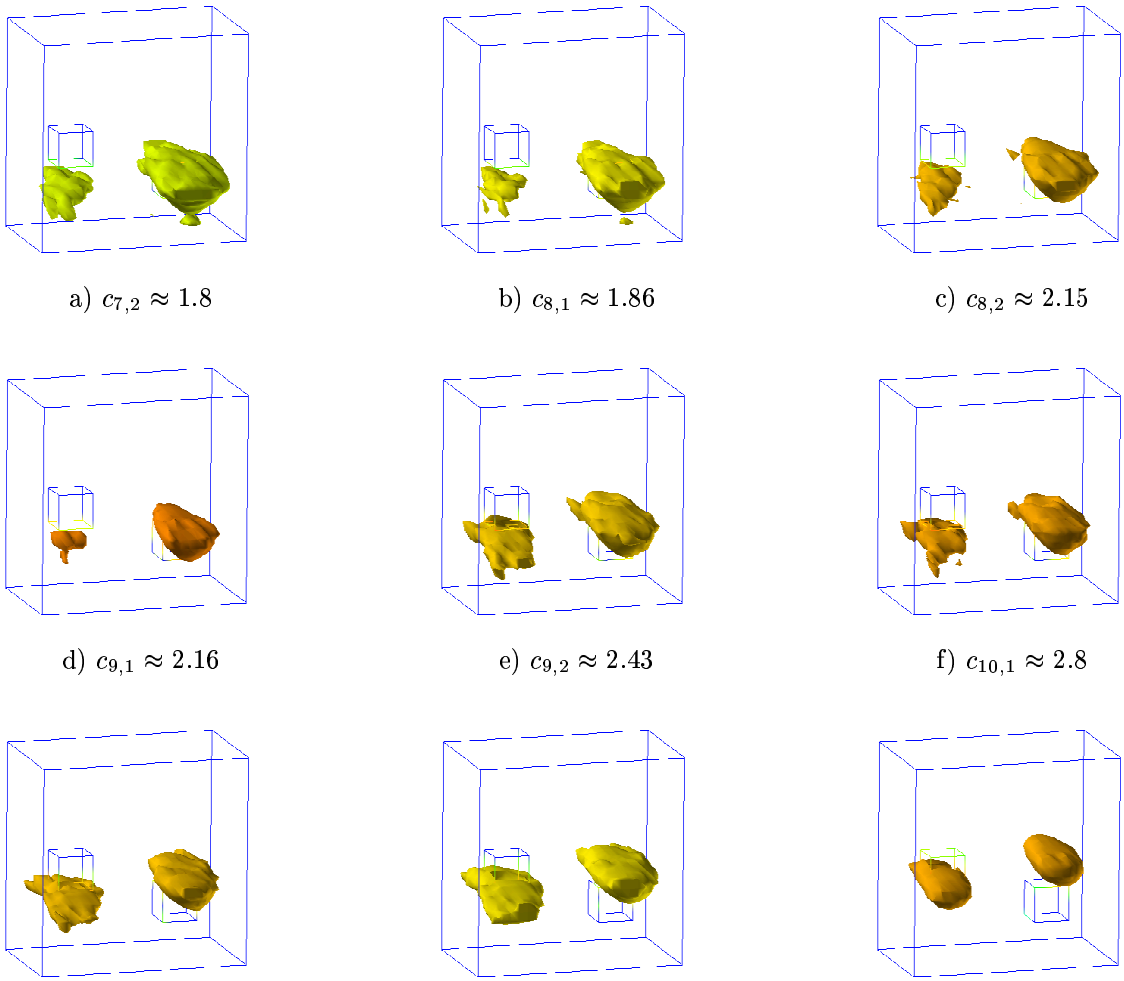
a) $c_{1,1} \approx 0.86$ b) $c_{2,1} \approx 1.00062$ c) $c_{2,2} \approx 1.002$

d) $c_{3,1} \approx 1.0424$ e) $c_{3,2} \approx 1.09$ f) $c_{4,1} \approx 1.125$

g) $c_{4,2} \approx 1.2$ h) $c_{5,1} \approx 1.26$ i) $c_{5,3} \approx 1.35$

j) $c_{6,1} \approx 1.4$ k) $c_{6,2} \approx 1.5$ l) $c_{7,1} \approx 1.6$

FIG. 9.3. *Test 1: the spatial distribution of c_h after computing $q_{n,k}; n = 1,...,7$ where n is the number of the computed function q. The staring point in adaptive algorithm is $c_{7,1}$. Results are presented with numerically approximated integrals $I_0, I_{1,n}, A_{1,n}, A_{2,n}$, with the noise level $\sigma = 5\%$, and with $\lambda = 200$.*



a) $c_{7,2} \approx 1.8$ b) $c_{8,1} \approx 1.86$ c) $c_{8,2} \approx 2.15$

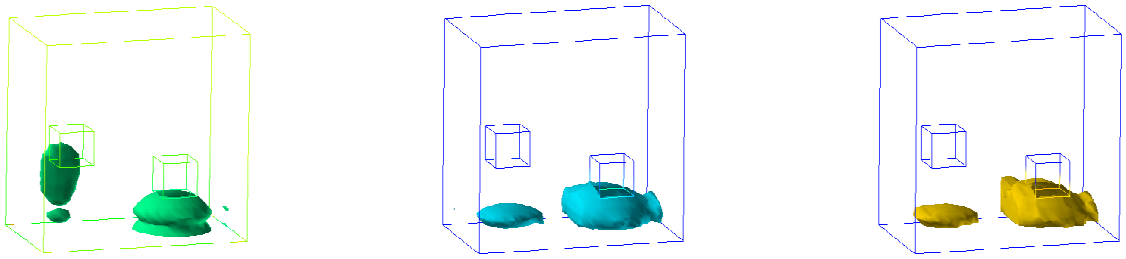
d) $c_{9,1} \approx 2.16$ e) $c_{9,2} \approx 2.43$ f) $c_{10,1} \approx 2.8$

g) $c_{11,1} \approx 3.28$ h) $c_{12,2} \approx 3.66$ i) $c_{13,2} \approx 4.09$

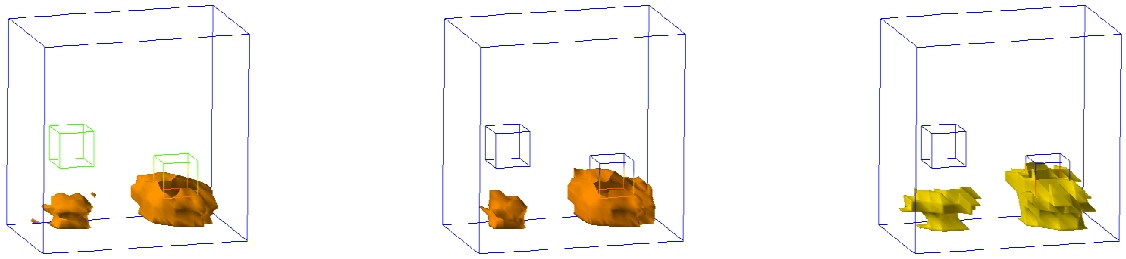
FIG. 9.4. *Test 1: the spatial distribution of c_h after computing $q_{n,k}; n = 7, ..., 13$ where n is the number of the computed function q. Results are presented with numerically approximated integrals $I_0, I_{1,n}, A_{1,n}, A_{2,n}$ and with $\sigma = 5\%, \lambda = 200$.*

where $\zeta = const \in (0, 1)$ is the tolerance number of our choice. The choice of the parameter ζ depends on the behaviour of the computed value of $\max |R_c(x)|$ in the right hand side of (9.3), see discussion in [6]. Thus, the choice of ζ depends on concrete values of the $|R_c(x)|$ and this should be done in numerical experiments. In (9.3) we take $\zeta = 0.2$ on the all computational meshes.

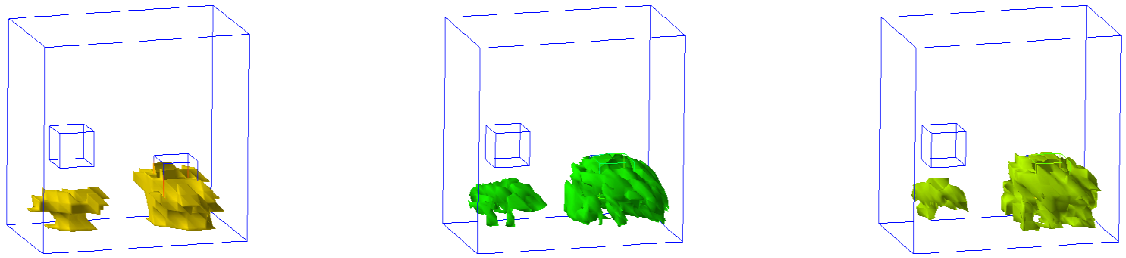
On all refined meshes we have used a cut-off parameter C_{cut} for the reconstructed coefficient



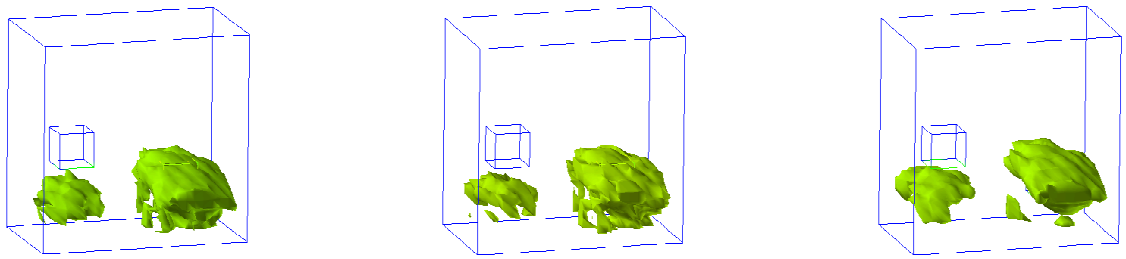
a) $c_{1,1} \approx 0.86$ b) $c_{3,1} \approx 1.0014$ c) $c_{4,1} \approx 1.042$



d) $c_{4,2} \approx 1.094$ e) $c_{5,1} \approx 1.136$ f) $c_{5,2} \approx 1.206$

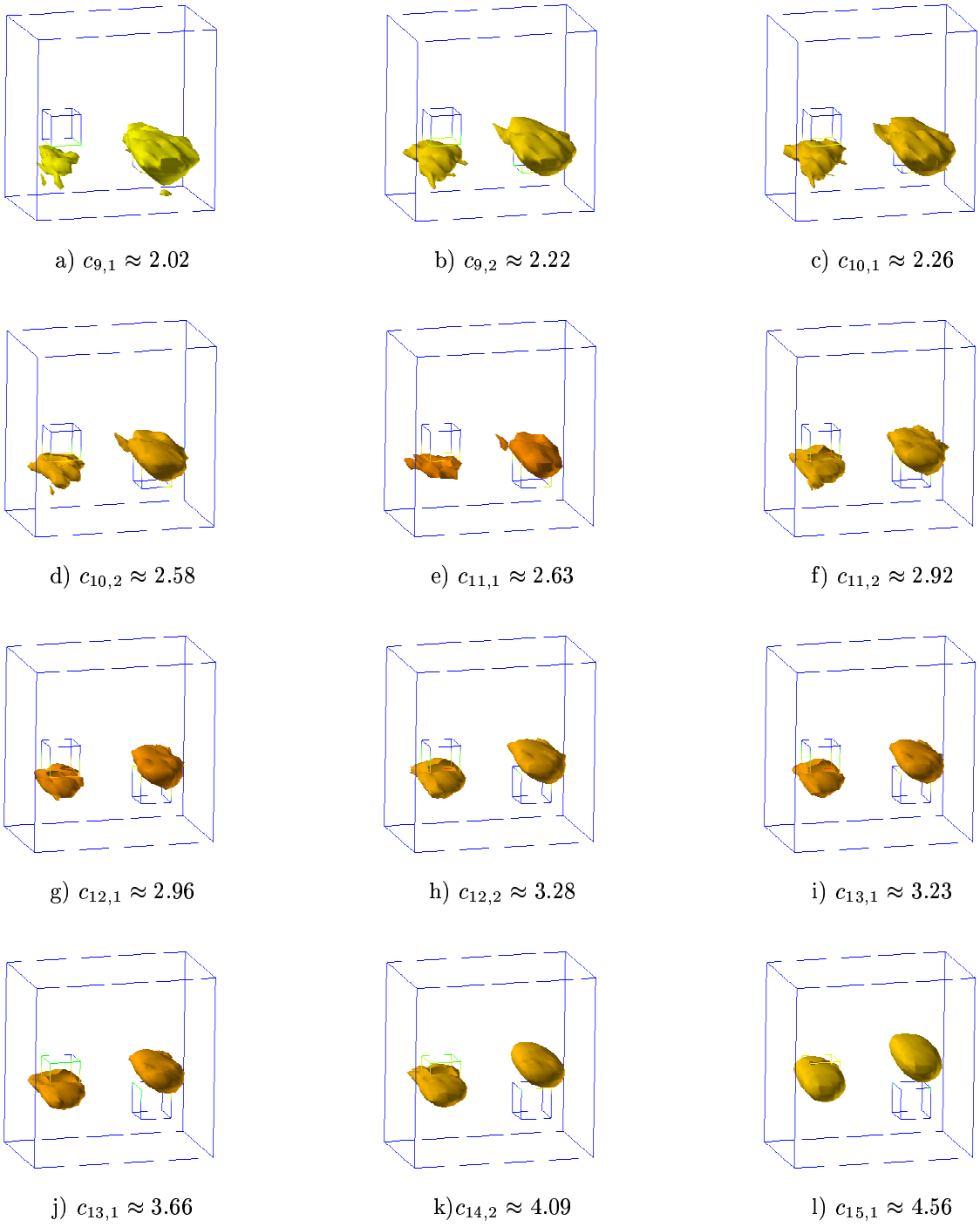


g) $c_{6,1} \approx 1.235$ h) $c_{6,2} \approx 1.384$ i) $c_{7,1} \approx 1.42$



j) $c_{7,2} \approx 1.5$ k) $c_{8,1} \approx 1.55$ l) $c_{8,2} \approx 1.68$

FIG. 9.5. *Test 1: the spatial distribution of c_h after computing $q_{n,k}; n = 1, ..., 8$ where n is the number of the computed function q. Results are presented with exactly approximated integrals $I_0, I_{1,n}, A_{1,n}, A_{2,n}$, with the noise level $\sigma = 5\%$, and with $\lambda = 200$.*



a) $c_{9,1} \approx 2.02$ b) $c_{9,2} \approx 2.22$ c) $c_{10,1} \approx 2.26$

d) $c_{10,2} \approx 2.58$ e) $c_{11,1} \approx 2.63$ f) $c_{11,2} \approx 2.92$

g) $c_{12,1} \approx 2.96$ h) $c_{12,2} \approx 3.28$ i) $c_{13,1} \approx 3.23$

j) $c_{13,1} \approx 3.66$ k)$c_{14,2} \approx 4.09$ l) $c_{15,1} \approx 4.56$

FIG. 9.6. *Test 1: the spatial distribution of c_h after computing $q_{n,k}; n = 9, ..., 15$ where n is the number of the computed function q. Results are presented with exactly approximated integrals $I_0, I_{1,n}, A_{1,n}, A_{2,n}$, with the noise level $\sigma = 5\%$, and with $\lambda = 200$.*

c_h such that

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

We choose $C_{cut} = 0.05$ for $m < 3$, $C_{cut} = 0.2$ for $m \geq 3$ in all tests. Here m is the number of iterations in quasi-Newton method. The application of the adaptivity technique allows us to get more correct locations of both small cubes depicted in Figure 7.1c.

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 Test 2 minimal and maximal values of the target coefficient in box constraints are taken using results of Test 1. To choose box constraints, we note first that we have already used the knowledge of the lower boundary $d_1 = 1$ for the function $c(x)$. To choose the upper bound, we observe that it follows from Figures 9.3-9.6 that before the above mentioned steep growth with n of norms $F_{n,i}$ for $n \geq 12$ the maximal value of the imaged coefficient did not exceed 4. Hence, we choose in (2.3) $d_2 = 2.1$. Thus, in Test 2 we enforce that the coefficient $c(x)$ belongs to the set of admissible parameters, $c(x) \in C_M = \{c \in C(\overline{\Omega}) | 1 \leq c(x) \leq 4.2\}$.

Testing was performed on 3-4 times adaptively refined meshes on 16 parallel processors. Results of computations are presented in Table 9.3. The relative time T_{rel} in these tables (CPU time/node) is computed as

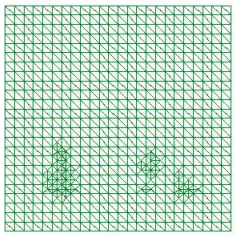
$$T_{rel} = \frac{T}{n}, \tag{9.4}$$

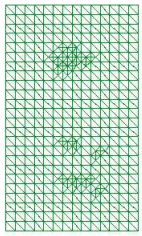
where T is the total CPU time, n is number of the nodes in computational mesh. We note that T_{rel} is approximately the same for all refined meshes which show efficiency of using hybrid FEM/FDM method for solution of the inverse problem. The knowledge of T_{rel} can help to estimate in advance the timing T for the solution of this CIP for any number of mesh points using (9.4).

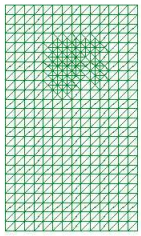
9.5. Test 2. This test actually consists of three tests with 0%, 3% and 5% noise, see beginning of subsection 9.4. In all three tests the function $c_{7,1}$, which corresponds to Figure 9.3-l, was taken as the starting point on the coarse mesh. In this test it was sufficient to use four adaptively refined meshes shown in Table 9.3.

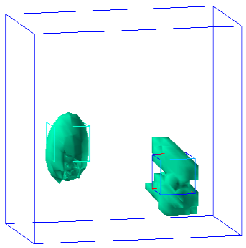
First, on the coarse mesh we get the same reconstruction as in globally convergent method, which is similar with the 2-D case of [6]. We refine mesh locally by computing residuals as in step 5 of the adaptive algorithm described in subsection 8.2. We repeat the optimization procedure on every new mesh. Already on the two times adaptively refined mesh we have reconstructed shifted location of cubes, see Figures 9.10. Table 9.3 shows computed L_2-norms of $||u|_{\Gamma_T} - g||_{L_2(\Gamma_T)}$ as also residuals R_{c_1} and R_{c_2}. We observe that residuals are decreasing as meshes are refined. Residuals are slightly increased for all refinements $n > 3$ of the initial mesh. All refinements $n > 3$ of the initial mesh does not improve results of the reconstruction. We conclude that on the three times adaptively refined mesh we get solution to our CIP, where we obtain the reconstructed coefficient $c_h \approx 4.0$. Thus, our final images for all three levels of noise are presented on Figures 9.7p, 9.8p and 9.9p.

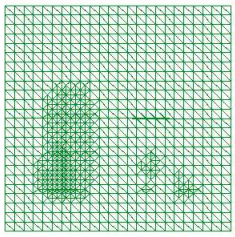
10. Summary. We have presented a further development of the new globally convergent numerical method of [5]. This method does not use a least squares residual functional. Instead, it relies on the structure of the underlying PDE operator. It was numerically demonstrated that

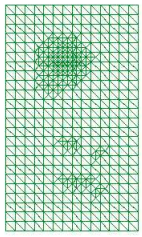

a) xy-projection

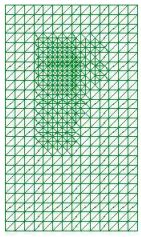

b) zx-projection

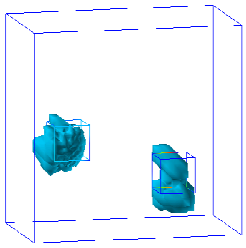

c) zy-projection

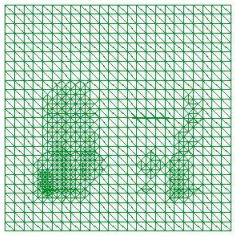

d) $c_h \approx 1.84$

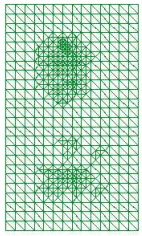

e) xy-projection

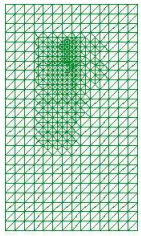

f) zx-projection

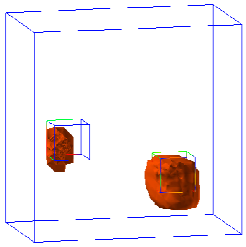

g) zy-projection

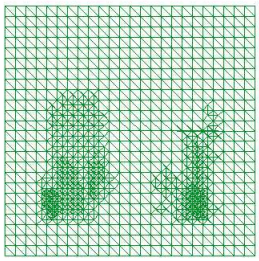

h) $c_h \approx 2.75$

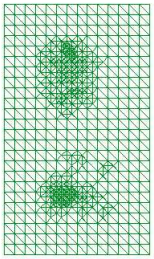

i) xy-projection

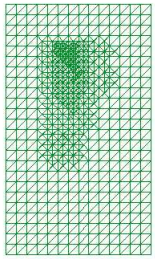

j) zx-projection


k) zy-projection


l) $c_h \approx 4.0$


m) xy-projection


n) zx-projection


o) zy-projection

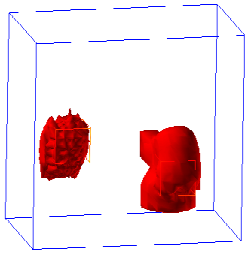
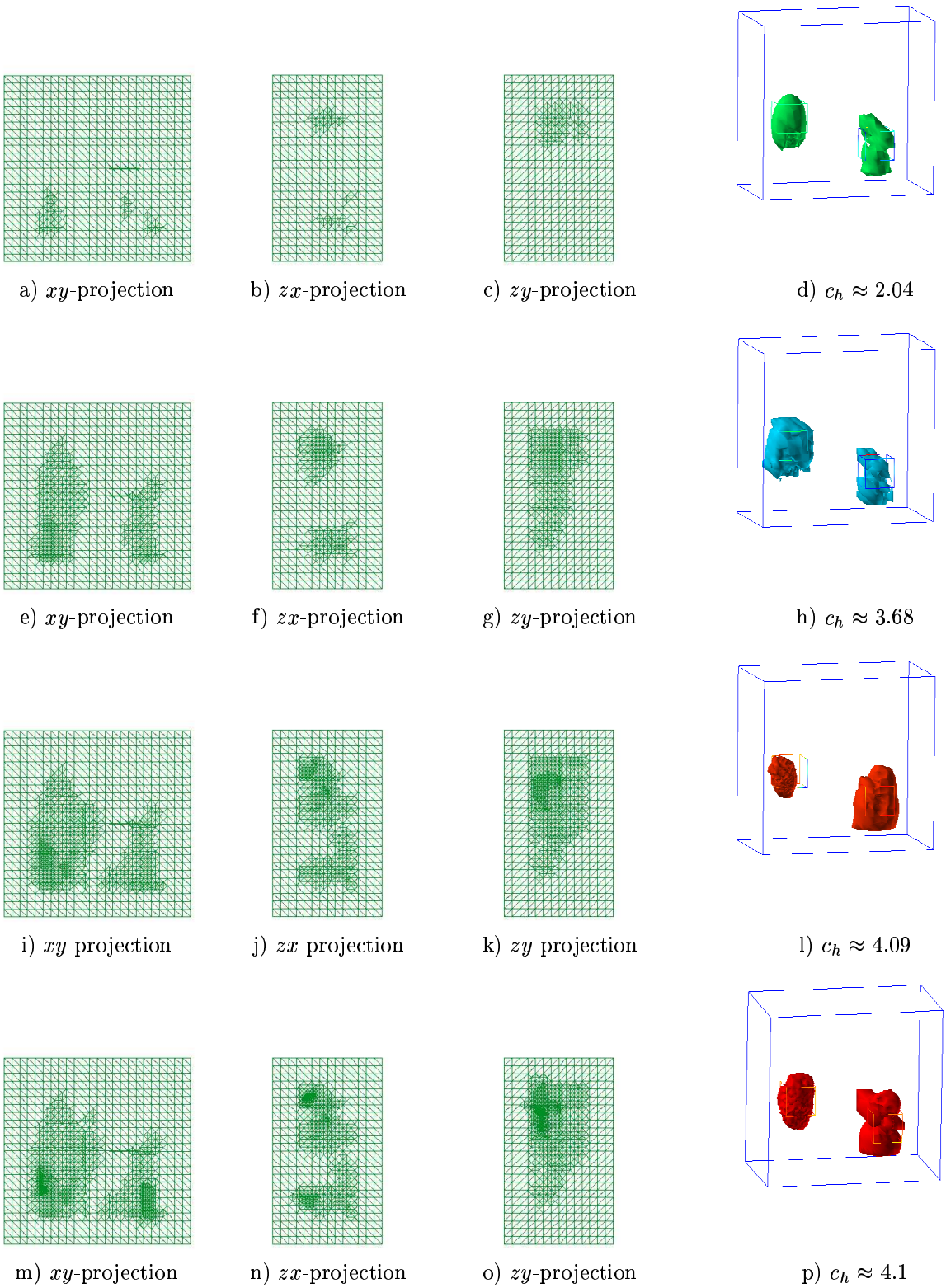

p) $c_h \approx 4.1$

FIG. 9.7. *Test 2: Adaptively refined computational meshes in different projections and reconstruction coefficient, correspondingly. Noise level is* $\sigma = 0\%$. *See Table 9.3 for the number of mesh points.*



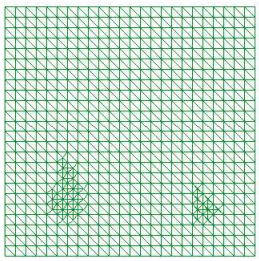
a) xy-projection b) zx-projection c) zy-projection d) $c_h \approx 2.04$

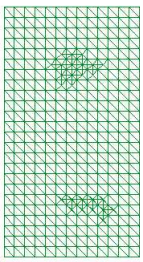
e) xy-projection f) zx-projection g) zy-projection h) $c_h \approx 3.68$

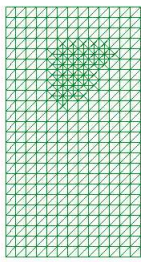
i) xy-projection j) zx-projection k) zy-projection l) $c_h \approx 4.09$

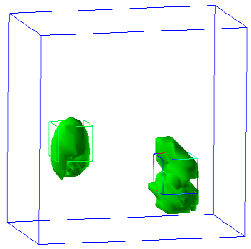
m) xy-projection n) zx-projection o) zy-projection p) $c_h \approx 4.1$

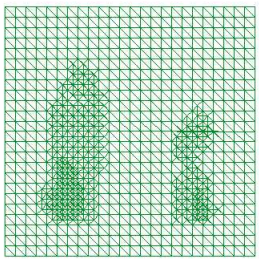
FIG. 9.8. *Test 2: Adaptively refined computational meshes in different projections and reconstruction coefficient, correspondingly. Noise level is* $\sigma = 3\%$. *See Table 9.3 for the number of mesh points.*

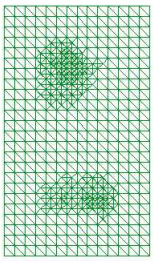

a) xy-projection

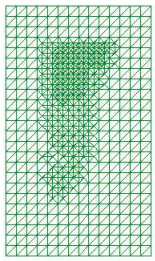

b) zx-projection

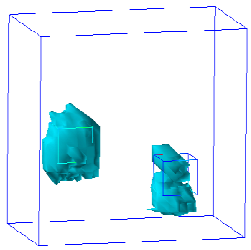

c) zy-projection

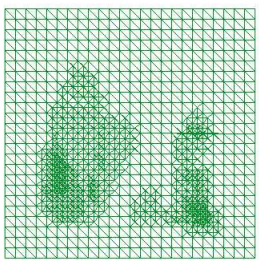

d) $c_h \approx 2.52$

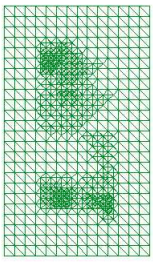

e) xy-projection

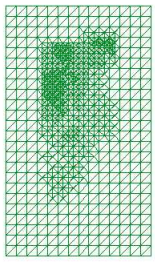

f) zx-projection

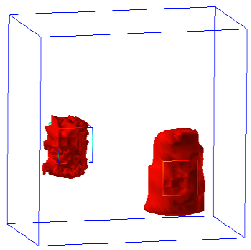

g) zy-projection

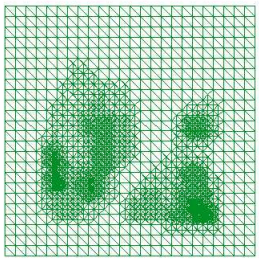

h) $c_h \approx 4.0$

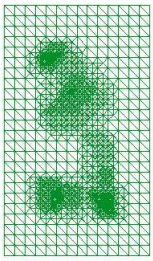

i) xy-projection

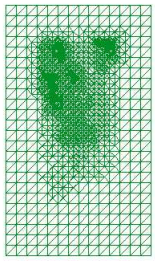

j) zx-projection


k) zy-projection


l) $c_h \approx 4.1$


m) xy-projection


n) zx-projection


o) zy-projection

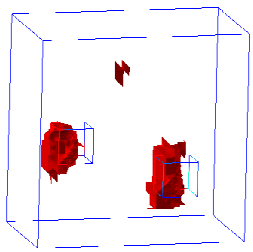
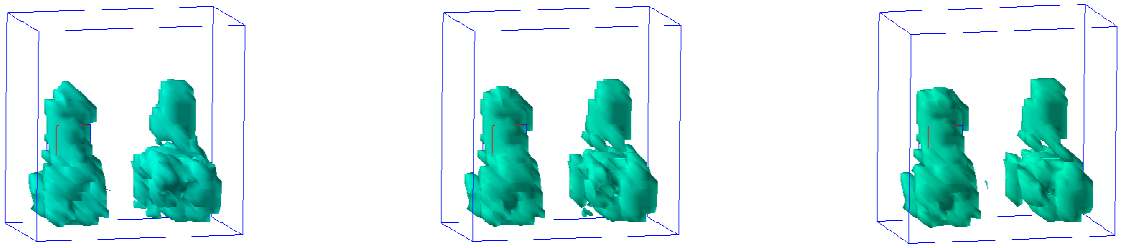

p) $c_h \approx 4.1$

FIG. 9.9. *Test 2: Adaptively refined computational meshes in different projections and reconstruction coefficient, correspondingly. Noise level is* $\sigma = 5\%$. *See Table 9.3 for the number of mesh points.*

Mesh $\sigma = 0\%$	$\|u\|_{\Gamma_T} - g\|_{L_2(\Gamma_T)}$	R_{c_1}	R_{c_2}	q.N.it.	CPU time (s) T	T_{rel}
9375	0.0285508	0.502511	0.0159757	5	23.87	0.0025
9583	0.0259143	0.358853	0.0440558	5	24.26	0.0025
10885	0.0301035	0.115057	0.105189	6	27.44	0.0025
11500	0.028857	0.119722	0.0952689	6	29	0.0025
12031	0.0342642	0.318698	0.049062	7	30.55	0.0025
Mesh $\sigma = 3\%$	$\|u\|_{\Gamma_T} - g\|_{L_2(\Gamma_T)}$	R_{c_1}	R_{c_2}	q.N.it.	CPU time (s) T	T_{rel}
9375	0.0293468	0.501782	0.0160071	5	23.88	0.0026
9569	0.0374917	0.21101	0.0427937	4	24.25	0.0025
11308	0.0299246	0.121369	0.0718892	6	28.36	0.0025
13153	0.0346342	0.118028	0.0784194	5	33.02	0.0025
14068	0.0365603	0.151333	0.0772201	4	35.09	0.0025
Mesh $\sigma = 5\%$	$\|u\|_{\Gamma_T} - g\|_{L_2(\Gamma_T)}$	R_{c_1}	R_{c_2}	q.N.it.	CPU time (s) T	T_{rel}
9375	0.031286	0.501337	0.0160262	4	23.77	0.0025
9555	0.0417805	0.18959	0.0497364	6	24.16	0.0025
11248	0.0293965	0.114448	0.0733725	6	28.18	0.0025
13042	0.0296054	0.126106	0.0723502	6	32.64	0.0025
20229	0.0398704	0.210689	0.105882	4	50.74	0.0025

TABLE 9.3

Test 2: $\|u\|_{\Gamma_T} - g\|_{L_2(\Gamma_T)}$, R_{c_1} *and* R_{c_2} *on adaptively refined meshes. Here, q.N.it. denotes the number of iterations in the quasi-Newton method. In this table coarse mesh consists of 9375 nodes. CPU time T is given for one q.N.it.*



a) $\sigma = 0\%$ b) $\sigma = 3\%$ c) $\sigma = 5\%$

FIG. 9.10. *Test 2: the spatial distribution of* c_h *on the coarse mesh with 9375 nodes and with different noise level* σ.

results obtained by the globally convergent part are in a good agreement with the convergence Theorem 6.1 as well as with the theory of ill-posed problems (subsection 9.3).

The convergence analysis implies that this method provides a good first guess for the Finite Element Adaptive method. This leads to a synthesis of both approaches. The main achievement of the adaptivity is that it provides the error analysis without a priori knowledge of the solution. When

the adaptivity is applied to an optimal control problem, it is a locally convergent numerical method. Hence, we have obtained a natural synthesis of both approaches. In this synthesis, the solution obtained by the globally convergent numerical method is used as a good first approximation for a further enhancement. Similarly with [6] we have observed that the application of the quasi-Newton method on the same mesh which was used by the globally convergent method, does not lead to an improvement of the solution. Hence, it is necessary to apply the adaptivity in order to refine mesh locally. Numerical results in the 3-D case demonstrate that the adaptivity enhances images obtained by the globally convergent method.

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, 2004.

[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] L. Beilina and M. V. Klibanov. A globally convergent numerical method and adaptivity for a hyperbolic coefficient inverse problem, *Chalmers Preprint Series* ISSN 1652-9715, 2009:6; also available online from http://www.ma.utexas.edu/mp_arc/.

[7] 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

[8] 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.

[9] 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.

[10] 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.

[11] 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/.

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

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

[14] 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.

[15] 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.

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

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

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

[19] 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.

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

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

[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. *International 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.