A posteriori error estimates for the adaptivity technique for the Tikhonov functional and global convergence for a coefficient inverse problem

LARISA BEILINA
MICHAEL V. KLIBANOV

A posteriori error estimates for the adaptivity technique for the Tikhonov functional and global convergence for a coefficient inverse problem

Larisa Beilina, Michael V. Klibanov

A posteriori error estimates for the adaptivity technique for the Tikhonov functional and global convergence for a coefficient inverse problem

Larisa Beilina*   Michael V. Klibanov †

December 18, 2009

Abstract

A synthesis of a globally convergent numerical method for a coefficient inverse problem and the adaptivity technique is presented. First, the globally convergent method provides a good approximation for the unknown coefficient. Next, this approximation is refined via the adaptivity technique. The analytical effort is focused on a posteriori error estimates for the adaptivity. A numerical test is presented.

1 Introduction

Coefficient Inverse Problems (CIPs) for PDEs are both nonlinear and ill-posed. These two factors cause substantial difficulties for their numerical treatments. This paper is a continuation of our previous publication [4], where a new globally convergent numerical method for a CIP for a hyperbolic PDE was developed. This CIP can be applied to inverse scattering of acoustical and electromagnetic waves. Compared with [4], the main new element here is that we combine the technique of [4] with the so-called Finite Element Adaptive method (adaptivity for brevity, see below in this section for the reason of this combination), see [5, 6] for earlier applications of the adaptivity to this specific CIP. Since the globally convergent numerical method was described in [4], we focus our analytical effort here on the adaptivity technique.

We present a new idea of obtaining a posteriori error estimates for the adaptivity for this CIP. This estimate is stronger than the one in [5, 6]. Specifically, certain integral terms, which were parts of a posteriori error estimates in [5, 6], were ignored in computational experiments there, because it was observed numerically that their absolute values were much smaller than

---

*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`).

the absolute value of a dominating term. An analytical explanation of this observation was not provided in [5, 6]. Unlike this, those integral terms are absent in our estimates, and the dominating term is the same as one in [5, 6], also see Remark 4.1 in section 4. In our numerical example the regularization parameter is chosen experimentally on the basis of the best performance. An analytical study of the question of the choice of the regularization parameter is outside of the scope of this publication. We refer to [17] for a detailed analysis of this interesting topic for the adaptivity technique.

We derive a posteriori error estimate for the Tikhonov functional. An alternative is to estimate the accuracy of the Lagrangian [3, 5, 6, 17]. The above new idea combines three elements: (1) derivation of Frechét derivatives of the so-called state and adjoint initial boundary value problems with respect to the unknown coefficient, (2) the follow up derivation of the Frechét derivative of the Tikhonov functional with the use of weak solution identities for state and adjoint problems, and (3) the Galerkin orthogonality [15]. We also present results of a numerical experiment.

This publication is driven by our numerical experience. When experimenting with the technique of [4], we have discovered that when we have to image a medium with either one inclusion or two inclusions located on the same horizontal level, as it is the case of [4], we always get a good image (the initializing plane wave propagates in the vertical direction). Next, when imaging two inclusions located on different horizontal levels, we saw that we can accurately image location of one of them as well as inclusions/background contrasts in both. At the same time, we have also observed that the location of the second inclusion is not imaged accurately by the technique of [4]. Namely, imaged inclusions remained on the same horizontal level, whereas it was desirable to move up one of them, e.g. compare Figures 1-c) and 3-c) in section 6. The global convergence estimate of [4] depends on a small parameter, which in turn depends on both the level of the error in the input data and approximation errors of the technique of [4]. While the error in the data cannot be avoided, approximation errors of the technique of [4] are not involved in locally convergent algorithms for CIPs. At the same time, although these errors are small, they cannot be made infinitely small in practical computations.

Therefore, our natural conclusion was that one should have a two-stage numerical procedure. On the first stage one should apply the technique of [4]. And on the second stage the solution obtained on the first should be refined via a subsequent application of an appropriate locally convergent numerical method. The latter method should take the solution of the globally convergent stage as a starting point for iterations. This is because a good first approximation is an important input which any locally convergent method needs. Our next question was about the choice of a proper locally convergent algorithm. When testing the quasi-Newton method on the same mesh where the globally convergent part worked, we have observed that the image quality was not improved: compare Figure 5-d) with Figure 3-c) in section 6. Thus, based on the previous experience of the first author [5, 6], we have chosen the adaptivity for the second stage. The adaptivity consists in applications of the quasi-Newton method on a sequence of adaptively refined meshes. So, the adaptivity helps indeed, see section 6.

2

One of the main ideas of the adaptivity is that for each mesh a posteriori error analysis shows subdomains where the biggest error of the computed solution is. Thus, an important point is that the mesh is refined *locally* in such subdomains. An alternative is to use a very fine mesh in the entire domain. However, the latter would lead to a substantial increase of both computing time and memory. Note that subdomains, where mesh is refined, are found without a priori knowledge of the solution. Instead one needs to know only an upper bound for the solution. In the case of classic forward problems, upper bounds are obtained from a priori estimates of solutions [1]. In the case of CIPs, upper bounds are assumed to be known in advance, which goes along well with the Tikhonov concept for ill-posed problems [13, 27]. A posteriori error analysis addresses the main question of the adaptivity: *Where to refine the mesh?* In the case of classic forward problems this analysis provides upper estimates for differences between computed and exact solutions locally, in subdomains of the original domain, see, e.g. [1, 15]. In the case of a forward problem, the main factor enabling to conduct a posteriori error analysis is the well-posedness of this problem. However, the ill-posed nature of CIPs *radically* changes the situation. It is because of the ill-posedness that an estimate of the difference between computed and exact coefficients is replaced by a posteriori estimate of the accuracy of either the Lagrangian [3, 5, 6, 17] or of the Tikhonov functional (the current paper).

In our numerical experiment we image a medium with small inclusions in it, although we do not assume a priori knowledge of such a structure. We refer to [2] and references cited there for another approach to imaging of small inclusions. In the globally convergent part we use a layer stripping procedure with respect to the pseudo frequency, which is the parameter $s > 0$ of the Laplace transform. A different layer stripping procedure with respect to the frequency was used in [11] for a numerical treatment of a CIP for the elliptic PDE, which can be obtained from the hyperbolic equation of this publication via the Fourier transform. Convergence theorem was not proven in this reference (see Remark 1.1 in [11]). There are also some other numerical methods for some multidimensional CIPs, which do not use a good first guess for the solution. While our technique works with the data resulting from a single measurement, these techniques work with the data resulting from multiple measurements [8, 9, 10, 18, 24, 25, 23], see [4] for a more detailed discussion of these works. In particular, the technique of [8, 9, 18] works with CIPs for hyperbolic PDEs.

This paper is organized as follows. In section 2 we state forward and inverse problems and outline the globally convergent numerical method of [4]. In section 3 we derive Frechét derivatives of solutions of the so-called state and adjoint initial boundary value problems as well as of the Tikhonov functional. In section 4 we establish a posteriori error estimates for the Tikhonov functional. In section 5 we outline the adaptive algorithm. In section 6 we present a computational example for the above two-stage numerical procedure.

3

2   Statements of forward and inverse problems and outline of the globally convergent technique of [4]

2.1   Statements of forward and inverse problems

As the forward problem, we consider the Cauchy problem for a hyperbolic PDE

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

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

Equation (1) governs 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, the dimensionless coefficient $c(x) = \varepsilon_r(x)$, where $\varepsilon_r(x)$ is the relative dielectric function of the medium, see [12], where this equation was derived from Maxwell's equations in the 2-D case. Let $\Omega \subset \mathbb{R}^3$ be a convex bounded domain with the boundary $\partial\Omega \in C^3$. We assume that the coefficient $c(x)$ of equation (1) is such that

$$c(x) \in [1, d], d = const. > 1, c(x) = 1 \text{ for } x \in \mathbb{R}^3 \backslash \Omega, \tag{3}$$

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

We consider the following

Inverse Problem. *Suppose that the coefficient $c(x)$ satisfies (3) and (4), where the number $d > 1$ is given. Assume that the function $c(x)$ is unknown in the domain Ω. Determine the function $c(x)$ for $x \in \Omega$, assuming that the following function $g(x, t)$ is known for a single source position $x_0 \notin \overline{\Omega}$*

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

A priori knowledge of upper and lower bounds of the coefficient $c(x)$ corresponds well with the Tikhonov concept about the availability of a priori information for an ill-posed problem [13, 27]. In applications the assumption $c(x) = 1$ for $x \in \mathbb{R}^3 \backslash \Omega$ means that the target coefficient $c(x)$ has a known constant value outside of the medium of interest Ω. Another argument here is that one should bound the coefficient $c(x)$ from the below by a positive number to ensure that the operator in (1) is a hyperbolic one on all iterations of our numerical procedure. Since we do not impose a "smallness" condition on the number $d - 1$, our numerical method is not a locally convergent one. The function $g(x, t)$ models time dependent measurements of the wave field at the boundary of the domain of interest. In practice measurements are performed at a number of detectors, of course. In this case the function $g(x, t)$ can be obtained via one of standard interpolation procedures, a discussion of which is outside of 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 (below) is approximately the same as one over $(0, \infty)$. The question of uniqueness of this Inverse Problem is a well known

long standing open problem. It is addressed positively only if the function $\delta(x - x_0)$ in (2) is replaced with a function $f(x)$ such that $f(x) \neq 0, \forall x \in \overline{\Omega}$. Corresponding uniqueness theorems were proven via the method of Carleman estimates [19, 20]. It is an opinion of the authors that because of applications, it makes sense to develop numerical methods, assuming that the question of uniqueness of the above inverse problem is addressed positively.

2.2 Outline of the technique of [4]

The outline of the globally convergent method of [4] in this subsection is given for the convenience of the reader. We refer to [4] for more details. Consider the Laplace transform of the functions u,

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

where $\underline{s}$ is a certain number. It is sufficient to choose $\underline{s}$ such that the integral (6) would converge together with corresponding (x, t)-derivatives. We call the parameter s *pseudo frequency*. Note that we do not use the inverse Laplace transform in our method, since approximations for the unknown coefficient are obtained in the pseudo frequency domain. We obtain from (1), (2)

$$\Delta w - s^2 c(x) w = -\delta(x - x_0), \tag{7}$$
$$\lim_{|x| \to \infty} w(x, s) = 0.$$

It follows from the classic theory of PDEs that for every $s > \underline{s}$ there exists unique solution $w \in C^{3+\gamma}(\mathbb{R}^3 \setminus \{|x - x_0| < \vartheta\}), \forall \vartheta > 0$. Here and below $C^{k+\gamma}, \gamma \in (0, 1)$ are Hölder spaces [21], where $k \geq 0$ is an integer. Since by the maximum principle $w(x, s) > 0$ [4], then we can consider the function $q(x, s) = \partial_s(s^{-2} \ln w(x, s))$. The function q satisfies the following nonlinear integral differential PDE with Volterra integrals with respect to s,

$$\Delta q - 2s^2 \nabla q \cdot \int_s^{\overline{s}} \nabla q(x, \tau) d\tau + 2s \left[\int_s^{\overline{s}} \nabla q(x, \tau) d\tau \right]^2$$
$$+ 2s^2 \nabla q \nabla V - 2s \nabla V \cdot \int_s^{\overline{s}} \nabla q(x, \tau) d\tau + 2s(\nabla V)^2 = 0, \tag{8}$$
$$q|_\Omega = \psi(x, s), \quad (x, s) \in \partial\Omega \times [\underline{s}, \overline{s}].$$

5

where the function ψ is generated by the function g in (5). The function $V(x, \bar{s})$ complements the rest of the integral, i.e.

$$
\frac{\ln w(x, s)}{s^2} = -\int_s^{\bar{s}} q(x, \tau)\, d\tau + V(x, \bar{s}),
$$

$$
V(x, \bar{s}) = \frac{\ln w(x, \bar{s})}{\bar{s}^2}. \tag{9}
$$

In (8), (9) $\bar{s}$ is the truncation pseudo frequency of integrals, which is one of regularization parameters of our method. Note that high frequencies are routinely truncated in science and engineering. We call $V(x, \bar{s})$ the "tail function", and it is unknown. It follows from (9) and Lemma 2.1 of [4] that the following asymptotic behavior is valid

$$
\|V(x, \bar{s})\|_{C^{2+\gamma}(\overline{\Omega})} = O\left(\frac{1}{\bar{s}}\right), \bar{s} \to \infty. \tag{10}
$$

We choose in our algorithm a sufficiently large value of $\bar{s}$ and assume, because of (10), that all approximations $V_{n,k}$ of the tail function, which are involved in our iterative procedure (below in this section), are bounded from the above by a small parameter ξ, i.e. $\|V_{n,k}\|_{C^{2+\gamma}(\overline{\Omega})} \leq \xi$ [4].

Equation (8) has two unknown functions, q and V. The reason why we can approximate well both of them is that we treat them differently. While we approximate the function q via inner iterations, the function V is approximated via outer iterations. The numerical solution of the problem (8) is the most challenging issue in this method, since this problem is nonlinear. The problem (8) is solved via a layer stripping procedure with respect to the pseudo frequency s. Consider a partition of the interval $[\underline{s}, \bar{s}]$ into N small subintervals of the width $\rho = s_{n-1} - s_n$, where $\underline{s} = s_N < s_{N-1} < ... < s_1 = \bar{s}$. Assume that $q(x, s)$ is a piecewise constant function with respect to s, $q(x, s) = q_n(x)$ for $x \in (s_n, s_{n-1}]$. Consider the Carleman Weight Function (CWF) $e^{\mu(s - s_{n-1})}$, where $\mu \gg 1$ is a large parameter which should be chosen in computations. Let in equation (8) $s \in (s_n, s_{n-1})$. We multiply this equation by the CWF function and integrate with respect to $s \in (s_n, s_{n-1})$. Hence, we obtain the Dirichlet boundary value problem for the following coupled system of nonlinear elliptic PDEs of the second order

$$
\begin{aligned}
L_n(q_n) &:= \Delta q_n - A_{1,n}\left(h\sum_{i=1}^{n-1}\nabla q_i\right)\nabla q_n + A_{1n}\nabla q_n \nabla V_n - \varkappa q_n \\
&= B_n(\nabla q_n)^2 - A_{2,n}h^2\left(\sum_{i=1}^{n-1}\nabla q_i(x)\right)^2 \\
&\quad + 2A_{2,n}\nabla V_n\left(h\sum_{i=1}^{n-1}\nabla q_i\right) - A_{2,n}(\nabla V_n)^2,
\end{aligned} \tag{11}
$$

$$
q_n\mid_{\partial\Omega} = \psi_n(x), n = 1, ..., N.
$$

Here the boundary data $\psi_n(x)$ are generated by the function $\psi(x, s)$ in (8), $A_{1,n}, A_{2,n}, B_n$ are certain numbers depending on μ, ρ, n and $\varkappa > 0$ is a small parameter of ones choice. This parameter is introduced to obtain a better stability of the problem (11) because of the maximum principle, see §1 in Chapter 3 of [21] . We write in (11) V_n instead of V for the convenience of the further description of this method. We have $\lim_{\mu \to \infty} B_n = 0$ uniformly for all n, due to the presence of the CWF. Hence, the presence of the CWF with $\mu >> 1$ mitigates the influence of the nonlinear term $(\nabla q_n)^2$, which enables us to solve the boundary value problem (11) for each q_n iteratively via solving a linear elliptic problem on each step. Still, the computational experience shows that we cannot take μ exceedingly large, which would effectively turn equations (11) into linear ones.

Problems (11) can be solved sequentially starting from $n = 1$, and this is exactly what we do. We use both inner and outer iterations to solve these problems. Let $\widetilde{w}(x, \overline{s})$ be the solution of the problem (7) with $c \equiv 1, s := \overline{s}$. Using (9), we set $V_{1,1}(x) := \overline{s}^{-2} \ln \widetilde{w}(x, \overline{s})$. So, first we find functions $q_1^k(x), k = 1, \dots$ via the iterative solution of the problem (11) for $n = 1$ with $V_1 := V_{1,1}$ and setting for the nonlinear term $(\nabla q_n)^2 := (\nabla q_1^{k-1})^2, q_1^0 = 0$. In other words, we iterate with respect to the nonlinear term until convergence occurs. The resulting function is denoted as $q_{1,1}$. For $n = 1$, we do not iterate with respect to the nonlinear term anymore, but rather iterate with respect to the tail as follows. Suppose that we have obtained the pair $(q_{1,k}, V_{1,k})$. Then we find the approximation $c_{1,k}(x)$ for the target coefficient $c(x)$ via backwards calculations. Next, we solve the problem (1), (2) with $c := c_{1,k}$, calculate the Laplace transform $w_{1,k+1}(x, s)$ of its solution and by (9) we set the new approximation for the tail as $V_{1,k+1}(x) := \overline{s}^{-2} w_{1,k+1}(x, \overline{s})$. Next, we solve the boundary value problem (11) with $V_1 := V_{1,k+1}, (\nabla q_n)^2 := (\nabla q_{1,k})^2$ and obtain the function $q_{1,k+1}$ this way. We continue this process until convergence occurs. Suppose that the convergence occurs at $k := m_1$. Then we set $(q_1, c_1, V_{2,1}) := (q_{1,m_1}, c_{1,m_1}, V_{1,m_1})$ and for the nonlinear term in (11) we set $(\nabla q_n)^2 := (\nabla q_2^0)^2 := (\nabla q_1)^2$. To find the function q_2, we repeat the above process for $n = 2$, etc., until convergence occurs at $k := m_2$, set $(q_2, c_2, V_{3,1}) := (q_{2,m_2}, c_{2,m_2}, V_{2,m_2}), (\nabla q_3^0)^2 := (\nabla q_2)^2$ and repeat for $n = 3$, etc.. So, for each n we iterate with respect to the nonlinear term only to approximate $q_{n,1}$. Numerical convergence criteria are specified in section 6, and so we stop computing functions c_n at $n := \overline{N} \in [1, N]$, where the final iteration number $\overline{N}$ is chosen on the basis of such a criterion, see (53) and (54). So, we denote $c_{glob}(x) := c_{\overline{N}}(x)$, and this is our solution obtained on the globally convergent stage. It is ensured that the function $c_{glob} \in Y$, where the set Y is defined in (14) in the next section. We refer to page 157 of [13], which explains that the number of iterations can serve as a regularization parameter for an ill-posed problem.

3   Frechét Derivatives

The first step of the adaptivity is the calculation of the Frechét derivative of the Tikhonov functional. To do this, we need, to calculate Frechét derivatives of state and adjoint initial boundary value problems. To achieve the latter, we need in turn to establish a certain

smoothness of solutions of state and adjoint initial boundary value problems. This smoothness cannot be guaranteed for the solution of the problem (1), (2) because of the $\delta-$ function in the initial condition. Hence, we assume in sections 3 and 4 that the δ-function in condition (2) is replaced with a regularized one,

$$u(x,0) = 0, u_t(x,0) = \delta_\theta(x - x_0),\tag{12}$$

where

$$\delta_\theta(x - x_0) = \left\{ \begin{array}{c} C_\theta \exp\left(\frac{1}{|x-x_0|^2-\theta^2}\right), \ |x - x_0| < \theta \\ 0, \ |x - x_0| \geq \theta \end{array} \right\}, \int\limits_{\mathbb{R}^m} \delta_\theta(x - x_0) \, dx = 1,\tag{13}$$

where $\theta > 0$ is so small that $\delta_\theta(x - x_0) = 0$ for $x \in \Omega$ (recall that $x_0 \neq \overline{\Omega}$) and the constant $C_\theta > 0$ is chosen to ensure the value of the integral in (13). Introduce the set Y of functions $c(x)$ satisfying the following conditions

$$Y = \left\{ \begin{array}{c} c \in C(\mathbb{R}^3), c - 1 \in H^1(\mathbb{R}^3), c(x) = 1 \text{ in } \mathbb{R}^3 \backslash \Omega \\ c_{x_i} \in L_\infty(\Omega), c(x) \in (1 - \omega, d + \omega) \text{ for } x \in \overline{\Omega} \end{array} \right\},\tag{14}$$

where $\omega \in (0, 1)$ is a small positive number. Let $T = const. > 0$. It follows from results of Chapter 4 of [22] that the solution of the problem (1), (12) $u \in C^\infty(\mathbb{R}^3 \times [0, T]), \forall c \in Y$.

Our CIP is a complex problem featuring both nonlinearity and ill-posedness combined with many yet unknown factors. Hence, it is not surprising that some discrepancies between our theories of two stages of our numerical procedure as well as between our theory and the computational practice take place. We now list these discrepancies. The replacement of (2) with (12), (13) is the first of these discrepancies. Also, conditions (14) are more general ones than (4), and this is the second discrepancy. The third discrepancy is that in our numerical study we replace the $\delta-$ function in (2) with the plane wave. This is due some conveniences of our numerical implementations in the past [4, 5, 6]. We point out, however that we use the $\delta-$ function in (2) only to obtain the asymptotic behavior (10), see Lemma 2.1 in [4]. In our numerical studies we verify this behavior computationally, see section 7.2 in [4]. The fourth discrepancy is that although we work with domains with sufficiently smooth boundaries in our analytical part, in the computational practice we use rectangles. Although, in principle at least, one might anticipate singularities of solutions near the corners of rectangles, we have not observed them in our computations.

3.1 State and adjoint problems and their Frechét derivatives

Let the function $c \in Y$. Since $c(x) = 1$ outside of the domain Ω, then, given the function g in (5), one can uniquely solve the initial boundary value problem (1), (5), (12) in the domain $(\mathbb{R}^3 \backslash \Omega) \times (0, T)$. Thus, we can uniquely find the function u in this domain. Let Ω_1 be a convex bounded domain such that $\Omega \subset \Omega_1, \partial\Omega \cap \partial\Omega_1 = \varnothing, \partial\Omega_1 \in C^\infty$ and $\delta_\theta(x - x_0) = 0$ in $\overline{\Omega}_1$. Denote $Q_T = \Omega_1 \times (0, T), S_T = \partial\Omega_1 \times (0, T)$. We assume that there exists a function

$a \in C^\infty\left(\overline{\Omega}_1\right)$ such that $a\mid_{\partial\Omega}= 0, \partial_n a\mid_{\partial\Omega}= 1$. For example, if $\Omega = \{|x| < R\}$, then one can choose $a\left(x\right) = \left(|x| - R\right)\chi\left(|x|\right)$, where the function χ is such that

$$\chi\left(z\right) \in C^\infty\left[0, R\right], \chi\left(z\right) = \left\{ \begin{array}{c} 1 \text{ for } z \in \left[\frac{R}{2}, R\right]; \\ 0 \text{ for } z \in \left[0, \frac{R}{4}\right]; \\ \text{between 0 and 1 for } z \in \left[\frac{R}{4}, \frac{R}{2}\right]. \end{array} \right\}.$$

Although such functions $a\left(x\right)$ might also likely be constructed for more general convex domains, we are not doing this here for brevity. Let $\widetilde{g}\left(x, t\right) = u\mid_{S_T}$, $p(x,t) = \partial_n u\mid_{S_T}$. Since the function u can be uniquely determined in $\left(\mathbb{R}^3\backslash\Omega\right) \times \left(0, T\right)$, then functions $\widetilde{g}, p$ can also be uniquely determined. It turns out that classic theorems about existence of solutions of initial boundary value problems for hyperbolic PDEs require that the boundary condition should have a sufficiently smooth extension inside the domain of interest, see, e.g., sections 4 and 5 in Chapter 4 of [22] as well as Theorems 5 and 6 in section 7.2 of [16]. Hence, we assume that there exist two functions F, W such that

$$\begin{aligned} F, W &\in& H^5\left(Q_T\right), &\quad (15) \\ \partial_n F\mid_{S_T} &=& p\left(x, t\right), \partial_n W\mid_{S_T}= \widetilde{g}\left(x, t\right), &\quad (16) \\ F\left(x, t\right) &=& W\left(x, t\right) = 0 \text{ for } x \in \Omega, &\quad (17) \\ \partial_t^j F\left(x, 0\right) &=& 0 \text{ in } \Omega_1, j = 0, ..., 3. &\quad (18) \end{aligned}$$

We impose these assumptions because the function g in (5) might be given with an error, meaning that the solution of the initial boundary value problem (1), (5), (12) would not necessarily belong to C^∞ then. So, better to consider rather generic functions $\widetilde{g}$ and p. If, however, the function g is given without an error, then we can set $F\left(x, t\right) = \chi_1\left(x\right)u, W\left(x, t\right) = \chi_1\left(x\right)a\left(x\right)u$, where the function $\chi_1 \in C^\infty\left(\overline{\Omega}_1\right)$ is such that $\chi_1\left(x\right) = 1$ near the boundary $\partial\Omega_1$ and $\chi_1\left(x\right) = 0$ in Ω. Existence of such functions χ_1 is known from the Real Analysis course. Consider now solutions u and λ of the following initial boundary value problems (we do not use a new notation for u for brevity),

$$\begin{aligned} c\left(x\right)u_{tt} &= \Delta u \text{ in } Q_T, \\ u\left(x, 0\right) &= u_t\left(x, 0\right) = 0, &\quad (19) \\ \partial_n u\mid_{S_T} &= p\left(x, t\right); \end{aligned}$$

$$\begin{aligned} c\left(x\right)\lambda_{tt} &= \Delta\lambda \text{ in } Q_T, \\ \lambda\left(x, T\right) &= \lambda_t\left(x, T\right) = 0, &\quad (20) \\ \partial_n\lambda\mid_{S_T} &= \left(\widetilde{g} - u\mid_{S_T}\right)z_\varepsilon\left(t\right). \end{aligned}$$

We call these problems the "state problem" and the "adjoint problem", respectively. Hence, it follows from (20) and (21) that for a given coefficient $c\left(x\right)$, one should first solve the state problem and next solve the adjoint problem. In (20) $z_\varepsilon\left(t\right)$ is a cut-off function, which is introduced to ensure that compatibility conditions at $\overline{S}_T \cap \{t = T\}$ are satisfied. Here $\varepsilon > 0$

9

is a small number. So, we choose such a function z_ε that

$$z_\varepsilon \in C^\infty [0, T], \ z_\varepsilon (t) = \left\{ \begin{array}{c} 1 \text{ for } \in [0, T - \varepsilon] \\ 0 \text{ for } t \in \left(T - \frac{\varepsilon}{2}, T\right] \\ \text{between 0 and 1 for } t \in \left(T - \varepsilon, T - \frac{\varepsilon}{2}\right) \end{array} \right\}.$$

We now reformulate for our specific needs a result, which follows immediately from Theorems 5 and 6 in section 7.2 of [16]. Although those theorems are proven for the Dirichlet boundary condition, modifications of those proofs to the case of the Neumann boundary condition are rather straightforward, see, e.g. Theorem 5.1 of Chapter 4 in [22]. Consider the following initial boundary value problem

$$\begin{aligned} c(x) v_{tt} &= \Delta v + f \text{ in } Q_T, \\ v(x, 0) &= v_t(x, 0) = 0, \\ \partial_n v \mid_{S_T} &= y(x, t) \in L_2(S_T), \end{aligned} \tag{21}$$

where the function $f \in H^k(Q_T), k \geq 0$. The weak solution $v \in H^1(Q_T)$ of this problem satisfies the following integral identity for all functions $r \in H^1(Q_T)$ with $r(x, T) = 0$

$$\int\limits_{Q_T} (-c(x) v_t r_t + \nabla v \nabla r) \, dx dt - \int\limits_{S_T} y r dx dt - \int\limits_{Q_T} f r dx dt = 0. \tag{22}$$

Assume that there exists such an extension $P(x, t)$ of the function $y(x, t)$ from the boundary S_T in the domain Q_T that $P \in H^{k+2}(Q_T), \partial_n P \mid_{S_T} = y(x, t), P(x, t) = 0$ for $x \in \Omega$, and in the case $k \geq 2$ let $\partial_t^j P(x, 0) = 0, j = 0, ..., k$ and $\partial_t^i f(x, 0) = 0, i = 0, ..., k - 2$. Consider the function $v - P$. Dividing both sides of equation (21) by $c(x)$ and using above cited theorems and the formula $c^{-1} \Delta v = \nabla \cdot (c^{-1} \nabla v) - \nabla (c^{-1}) \nabla v$, we obtain that actually the weak solution $v \in H^{k+1}(Q_T)$ and the following estimate holds

$$\|v\|_{H^{k+1}(Q_T)} \leq B \left[\|P\|_{H^{k+2}(Q_T)} + \|f\|_{H^k(Q_T)}\right]. \tag{23}$$

Here and below $B = B(Y, Q_T, a(x))$ and $C = C\left(B, z_\varepsilon, \|F\|_{H^5(Q_T)}, \|W\|_{H^5(Q_T)}\right)$ are different positive constants depending on listed parameters. Consider functions $\widehat{u} = u - F, \widehat{\lambda} = \lambda - (W - a(x) u) z_\varepsilon$ and substitute them in (19), (20). Then, using (15)-(18), (21) and (23), we obtain that functions $u, \lambda \in H^4(Q_T)$ and

$$\|u\|_{H^4(Q_T)} \leq B \|F\|_{H^5(Q_T)}, \ \|\lambda\|_{H^4(Q_T)} \leq B \left(\|F\|_{H^5(Q_T)} + \|W\|_{H^5(Q_T)}\right). \tag{24}$$

Introduce the set Z of functions defined in Ω_1,

$$Z = \left\{ f : f \in C\left(\overline{\Omega}_1\right) \cap H^1(\Omega_1), \partial_{x_i} f \in L_\infty(\Omega_1) \right\}.$$

10

Define the norm in Z as

$$\|f\|_Z := \|f\|_{C(\overline{\Omega}_1)} + \sum_{i=1}^{3} \|\partial_{x_i} f\|_{L_\infty(\Omega_1)}. \tag{25}$$

Then Z is a Banach space, since convergence in the norm $\|\cdot\|_Z$ implies convergence in both spaces $C(\overline{\Omega}_1)$ and $H^1(\Omega_1)$. Let $\widetilde{Y}$ be the set of restrictions of all functions of the set Y on the domain Ω_1. Then it follows from (14) and (25) that $\widetilde{Y}$ is an open set in the space Z and

$$c_1(x) - c_2(x) \in Z' := \{f \in Z : f(x) = 0 \text{ in } \Omega_1 \backslash \Omega\}, \quad \forall c_1, c_2 \in \widetilde{Y}. \tag{26}$$

In Theorem 3.1 we establish Frechét derivatives of state and adjoint problems with respect to the coefficient $c(x)$.

Theorem 3.1. *Assume that initial conditions (2) are replaced with initial conditions (12), where the function $\delta_\theta(x - x_0)$ is defined in (13). Let domains Ω, Ω_1 and the function $a(x)$ be those specified above and $\delta_\theta(x - x_0) = 0$ in Ω_1. Assume that there exist functions F, W satisfying conditions (15)-(18). Consider the set $\widetilde{Y}$ as an open set in the space Z. Let operators $A_1 : \widetilde{Y} \to H^1(Q_T)$ and $A_2 : \widetilde{Y} \to H^1(Q_T)$ map every function $c \in \widetilde{Y}$ in the weak solution $u(x, t, c)$ of the problem (19) and the weak solution $\lambda(x, t, c)$ of the problem (20) respectively, where in (20) $u\mid_{S_T} := u(x, t, c)\mid_{S_T}$. Then in fact functions $u(x, t, c), \lambda(x, t, c) \in H^4(Q_T)$ and each of the operators A_1 and A_2 has the Frechét derivative $A_1'(c)(b) = \widetilde{u}(x, t, c, b) \in H^1(Q_T)$ and $A_2'(c)(b) = \widetilde{\lambda}(x, t, c, b) \in H^1(Q_T)$ at each point $c \in \widetilde{Y}$, where $b(x) \in Z'$ is an arbitrary function. In fact, functions $\widetilde{u}, \widetilde{\lambda} \in H^2(Q_T)$ and they are solutions of the following initial boundary value problems*

$$
\begin{aligned}
&c(x)\widetilde{u}_{tt} = \Delta\widetilde{u} - b(x)u_{tt}(x, t, c), \text{ in } Q_T, \\
&\widetilde{u}(x, 0) = \widetilde{u}_t(x, 0) = 0, \partial_n\widetilde{u}\mid_{S_T} = 0;
\end{aligned} \tag{27}
$$

$$
\begin{aligned}
&c(x)\widetilde{\lambda}_{tt} = \Delta\widetilde{\lambda} - b(x)\lambda_{tt}(x, t, c), \text{ in } Q_T, \\
&\widetilde{\lambda}(x, T) = \widetilde{\lambda}_t(x, T) = 0, \partial_n\widetilde{\lambda}\mid_{S_T} = -z_\varepsilon\widetilde{u}\mid_{S_T}.
\end{aligned} \tag{28}
$$

Denote

$$A_3(c)(x) := \int_0^T (u_t\lambda_t)(x, t, c)\, dt, x \in \Omega, \forall c \in \widetilde{Y}.$$

Then the operator $A_3 : \widetilde{Y} \to C(\overline{\Omega})$.

Proof. The validity of the statement about the smoothness of functions u, λ follows from (24). Consider an arbitrary function $c \in \widetilde{Y}$. It follows from (14) that there exists a sufficiently small number $\varepsilon_1 \in (0, 1)$ such that $1 - \omega(1 - \varepsilon_1) \leq c(x) \leq d + \omega(1 - \varepsilon_1)$. Let the function $b \in Z'$ be such that $\|b\|_{C(\overline{\Omega}_1)} < \varepsilon_1\omega$, where the set Z' is defined in (26). Then $c + b \in \widetilde{Y}$. By (23)-(27) the function $\widetilde{u} \in H^2(Q_T)$ and

$$\|\widetilde{u}\|_{H^2(Q_T)} \leq B\|F\|_{H^5(Q_T)} \cdot \|b\|_Z. \tag{29}$$

11

Denote

$$u^{c+b}(x,t) \quad : \quad = u(x,t,c+b), u^c(x,t) := u(x,t,c),$$
$$u_1 \quad : \quad = u_1(x,t,c,b) = \left(u^{c+b} - u^c - \widetilde{u}\right)(x,t).$$

Hence, $u_1 \in H^2(Q_T)$. We now figure out the equation for the function u_1. By (19) and (27)

$$\begin{aligned}
\Delta u_1 &= (c+b) u_{tt}^{c+b} - cu_{tt}^c - c\widetilde{u}_{tt} - bu_{tt}^c = (c+b) u_{tt}^{c+b} - (c+b) u_{tt}^c - c\widetilde{u}_{tt} \\
&= (c+b)\left(u^{c+b} - u^c - \widetilde{u}\right)_{tt} + b\widetilde{u}_{tt} = (c+b) u_{1tt} + b\widetilde{u}_{tt}.
\end{aligned}$$

Hence, the function u_1 is the solution of the following intial boundary value problem

$$(c+b) u_{1tt} = \Delta u_1 - b\widetilde{u}_{tt}; \quad u_1(x,0) = u_{1t}(x,0) = 0, \quad \partial_n u_1 \mid_{S_T} = 0. \tag{30}$$

Hence, (23), (29) and (30) imply that

$$\|u_1\|_{H^1(Q_T)} \le C \|b\|_Z^2. \tag{31}$$

Hence,

$$\lim_{\|b\|_Z \to 0} \left(\frac{\|u_1\|_{H^1(Q_T)}}{\|b\|_Z} \right) = \lim_{\|b\|_Z \to 0} \left(\frac{\|u(x,t,c+b) - u(x,t,c) - \widetilde{u}(x,t,c,b)\|_{H^1(Q_T)}}{\|b\|_Z} \right) = 0. \tag{32}$$

Note that we set $A_i : \widetilde{Y} \to H^1(Q_T), i = 1,2$ instead of $A_i : \widetilde{Y} \to H^2(Q_T)$ only for the sake of the estimate (31). Since the function $\widetilde{u}(x,t,c,b)$ depends linearly on b, then (31) and (32) imply that the function $\widetilde{u}$ is indeed the Frechét derivative of the operator A_1 at the point c. Hence, we now can consider the function $\widetilde{u}(x,t,c,b)$ for any $b \in Z'$. The proof for the operator A_2 is similar. Finally, it follows from (24) and the embedding theorem that functions $u, \lambda \in C^1\left(\overline{Q}_T\right)$, which implies the statement about the operator A_3. $\square$

3.2   The Frechét derivative of the Tikhonov functional

Assume that conditions of Theorem 3.1 hold. We define the Tikhonov functional $E : \widetilde{Y} \to \mathbb{R}$ as

$$E(c) = \frac{1}{2} \int\limits_{S_T} (u \mid_{S_T} - \widetilde{g}(x,t))^2 z_\varepsilon(t) \, dxdt + \frac{1}{2}\alpha \int\limits_\Omega (c - c_{glob})^2 \; dx, \forall c \in \widetilde{Y},$$

where $\alpha \in (0,1)$ is the regularization parameter and $c_{glob} \in \widetilde{Y}$ is the approximation for the exact solution c^* obtained on the globally convergent stage, see the end of section 2. We use the domain Ω rather than Ω_1 in the second integral term because of (26). Consider the associated Lagrange functional $L(c)$,

$$L(c) = E(c) + \int\limits_{Q_T} \left(-c(x) u_t \lambda_t + \nabla u \nabla \lambda\right) dxdt - \int\limits_{S_T} p\lambda dxdt,$$
$$u := u(x,t,c) \in H^4(Q_T), \lambda := \lambda(x,t,c) \in H^4(Q_T), \tag{33}$$

12

where functions $u(x, t, c)$ and $\lambda(x, t, c)$ are solutions of initial boundary value problems (19), (20). The reason why we consider $L(c)$ is that we want to simplify the calculation of the Frechét derivative of $E(c)$. By (19), (20) and (22) the integral term in the first line of (33) equals zero. Hence,

$$L(c) = E(c) \text{ implying that } L'(c) = E'(c), \forall c \in \widetilde{Y},$$

where $L'(c)$ and $E'(c)$ are Frechét derivatives of functionals $L(c)$ and $E(c)$ respectively. To obtain the explicit expression for $L'(c)$, we need to vary in (33) the function c via considering $c + b \in \widetilde{Y}$ for $b \in Z'$ and then to single out the term, which is linear with respect to b. When varying c, we also need to consider respective variations of functions u and λ in (33), since these functions depend on c as solutions of state and adjoint problems (19) and (20). By Theorem 3.1, linear, with respect to c, parts of variations of u and λ are functions $\widetilde{u}(x, t, c, b)$, $\widetilde{\lambda}(x, t, c, b)$.

Theorem 3.2. *Assume that conditions of Theorem 3.1 hold. Then for every function* $c \in \widetilde{Y}$

$$E'(c) = L'(c) = \alpha(c - c_{glob}) - \int_0^T u_t \lambda_t dt, \tag{34}$$

$$E'(c) \in C\left(\overline{\Omega}\right).$$

Proof. Considering in (33) $L(c + b) - L(c) = E(c + b) - E(c)$, singling out the term, which is linear with respect to b, and using Theorem 3.1, we obtain

$$
\begin{aligned}
L'(c)(b) = E'(c)(b) = \int_{\Omega} &\left[\alpha(c - c_{glob}) - \int_0^T u_t \lambda_t dt \right] b(x) \, dx \\
&+ \int_{Q_T} \left(-c u_t \widetilde{\lambda}_t + \nabla u \nabla \widetilde{\lambda} \right) dx dt - \int_{S_T} p \widetilde{\lambda} dx dt \\
&+ \int_{Q_T} \left(-c \lambda_t \widetilde{u}_t + \nabla \lambda \nabla \widetilde{u} \right) dx dt - \int_{S_T} \left(\widetilde{g} - u \mid_{S_T} \right) z_{\varepsilon_1}(t) \widetilde{u} dx dt, \forall c \in \widetilde{Y}, \forall b \in Z',
\end{aligned}
\tag{35}
$$

where $\widetilde{u}$ and $\widetilde{\lambda}$ are solutions of problems (27) and (28) respectively. Since $\widetilde{u}(x, 0) = \widetilde{\lambda}(x, T) = 0$, then (19), (20) and (22) imply that second and third lines in (35) equal zero, which proves the first line of (34). The validity of the second line of (34) follows from the statement of Theorem 3.1 about the operator A_3. $\square$

4   A Posteriori Error Estimates in The Adaptivity

We are now ready to derive a posteriori error estimates for the Tikhonov functional $E(c)$ in the adaptivity technique. We work only with piecewise linear finite elements, because they

13

are used in our computations. Extensions on other finite elements are outside of the scope of this publication. Consider a finite element mesh with the maximal grid step size h. Let the function $f \in Z$ and let f^I be its standard interpolant on this mesh [15]. It follows from the formula 76.3 of [15] that

$$\left\| f - f^I \right\|_{C(\overline{\Omega}_1)} \leq K \left\| \nabla f \right\|_{L_\infty(\Omega_1)} h, \tag{36}$$

where the positive constant $K = K(\Omega_1)$ depends only on the domain Ω_1. Since we will work with finite element approximations of the target coefficient c, we introduce the space of finite elements C_h with the norm $\| \cdot \|_{C_h} := \| \cdot \|_{C(\overline{\Omega})}$. Since $\dim C_h < \infty$, then all norms in this space are equivalent. Also, $C_h \subset Z$ as a set. Hence, if the function $\widetilde{c}(x)$ is defined in Ω_1 and is such that

$$\begin{aligned} \widetilde{c}(x) \in C_h \text{ for } x \in \Omega_1; \ \ \widetilde{c}(x) \in (1 - \omega, d + \omega) \text{ in } \Omega; \\ \widetilde{c}(x) = 1 \text{ in } \Omega_1 \backslash \Omega, \text{ then } \widetilde{c}(x) \in \widetilde{Y}. \end{aligned} \tag{37}$$

Recall that the function $c^*(x)$ is the exact solution of our CIP and c^* satisfies (3), (4). Hence, it follows from (3), (4) that the restriction of the function c^* on the domain Ω_1 is such that $c^* \in \widetilde{Y}$ and the same is true for the restriction of the function c_{glob} on the domain Ω_1, i.e. $c_{glob} \in \widetilde{Y}$ (we do not change notation for brevity). Since the second stage of our two-stage procedure, the adaptivity, is a locally convergent numerical method, which the function c_{glob} as the starting point, we work in this section in a small neighborhood V_r of the exact solution c^*, where $V_r = \{c \in Z : \|c - c^\star\|_Z < r\}$, and $r > 0$ is a small number. It can be easily derived from (31) and Theorem 3.2 that

$$E(c_1) - E(c_2) = E'(c_2)(c_1 - c_2) + f(c_1, c_2), \forall c_1, c_2 \in V_r, \tag{38}$$

where $E'(c_2)$ is given in (34) and $f(c_1, c_2) = O(r^2), r \to 0$, uniformly for all $c_1, c_2 \in V_r$. Hence, in a posteriori error estimates for the adaptivity below we estimate only the accuracy of the calculation of the Frechét derivative of the Tikhonov functional. Thus, these will be approximate error estimates, since we will ignore the function f. The computational experience of section 6 as well as of the past publications [5, 6] shows that this approximation is sufficient.

In Theorem 4.1 we assume that state and adjoint problems are solved exactly for the case when the coefficient belongs to C_h.

Theorem 4.1. *Assume that conditions of Theorem 3.1 hold. Suppose that there exists a minimizer $c_\alpha \in \widetilde{Y}$ of the functional $E(c)$ on the set V_r as well as a minimizer $c_{\alpha h} \in \widetilde{Y}$ of $E(c)$ on the set $V_r \cap C_h$ (also, see (37)). Assume also that state and adjoint problems (19) and (20) are solved exactly for both coefficients c_α and $c_{\alpha h}$. Then the following approximate error estimate for the above defined Tikhonov functional is valid*

$$|E(c_\alpha) - E(c_{h\alpha})| \leq \left(A(\Omega) K \left\| E'(c_{\alpha h}) \right\|_{C(\overline{\Omega})} \right) \left\| \nabla c_\alpha \right\|_{L_\infty(\Omega)} h,$$

14

where $A(\Omega)$ is the volume of the domain Ω, K is the interpolation constant from (36) and by (34)

$$E'(c_{\alpha h}) = \alpha(c_{\alpha h} - c_{glob}) - \int_0^T (u_t \lambda_t)(x, t, c_{\alpha h}) \, dt, \tag{39}$$

where functions $u(x, t, c_{\alpha h}) \in H^4(Q_T)$ and $\lambda(x, t, c_{\alpha h}) \in H^4(Q_T)$ are solutions of problems (19) and (20) respectively with $c := c_{\alpha h}$.

Proof. Since the function $c_{h\alpha}$ is a minimizer of $E(c)$ on $V_r \cap C_h$, then

$$E'(c_{\alpha h})(b) = 0, \ \forall b \in C_h. \tag{40}$$

Now we use the Galerkin orthogonality [15]. We have splitting [15],

$$c_\alpha - c_{h\alpha} = (c_\alpha - c_\alpha^I) + (c_\alpha^I - c_{h\alpha}).$$

Since, $c_\alpha^I - c_{h\alpha} \in C_h$, then by (40) $E'(c_{\alpha h})(c_\alpha^I - c_{h\alpha}) = 0$. Hence,

$$E'(c_{\alpha h})(c_\alpha - c_{h\alpha}) = E'(c_{\alpha h})(c_\alpha - c_\alpha^I). \tag{41}$$

By Theorem 3.2 the function $E'(c_{\alpha h}) \in C(\overline{\Omega})$. Using (36), (38), (39) and (41), we obtain the following approximate error estimate

$$\begin{aligned}
|E(c_\alpha) - E(c_{h\alpha})| &\leq |E'(c_{\alpha h})(c_\alpha - c_\alpha^I)| \\
&\leq \|c_\alpha - c_\alpha^I\|_{C(\overline{\Omega})} \int_\Omega \left| \alpha(c_{\alpha h} - c_{glob}) - \int_0^T (u_t \lambda_t)(x, t, c_{\alpha h}) \, dt \right| dx \\
&\leq \left(A(\Omega) K \|E'(c_{\alpha h})\|_{C(\overline{\Omega})} \right) \|\nabla c_\alpha\|_{L_\infty(\Omega)} h. \quad \square
\end{aligned} \tag{42}$$

While it was assumed in Theorem 4.1 that state and adjoint problems (19) and (20) are solved exactly, in the computational practice they are solved approximately with a small error. Hence, it is desirable to express a posteriori error estimate through these approximate solutions. This is done in Theorem 4.2.

Theorem 4.2. *Assume that conditions of Theorem 3.1 hold. Suppose that there exists a minimizer $c_\alpha \in \widetilde{Y}$ of the functional $E(c)$ on the set V_r as well as a minimizer $c_{\alpha h} \in \widetilde{Y}$ of $E(c)$ on the set $V_r \cap C_h$. Suppose that state and adjoint problems (19) and (20) are solved exactly for $c := c_\alpha$ and that they are solved computationally with an error for $c := c_{\alpha h}$. Let functions $u_h := u_h(x, t, c_{\alpha h}) \in H^1(Q_T), \lambda_h := \lambda_h(x, t, c_{\alpha h}) \in H^1(Q_T)$ be those approximate solutions. Suppose that functions $u_{ht}, \lambda_{ht} \in L_\infty(Q_T)$. Let functions $u := u(x, t, c_{\alpha h}), \lambda := \lambda(x, t, c_{\alpha h}) \in H^4(Q_T)$ (see (24)) be exact solutions of problems (19) and (20) with $c := c_{\alpha h}$. Assume that*

$$\|u - u_h\|_{H^1(Q_T)} + \|\lambda - \lambda_h\|_{H^1(Q_T)} \leq \zeta, \tag{43}$$

where $\zeta \in (0,1)$ is a small number. Then the following approximate a posteriori error estimate is valid

$$|E(c_\alpha) - E(c_{h\alpha})| \leq K\left(A(\Omega)\|D(c_{\alpha h})\|_{L_\infty(\Omega)} + C\zeta\right)\|\nabla c_\alpha\|_{L_\infty(\Omega)} h, \qquad (44)$$

where the positive constant C was introduced in section 3 and

$$D(c_{\alpha h}) := \alpha(c_{\alpha h} - c_{glob}) - \int_0^T (u_{ht}\lambda_{ht})(x, t, c_{\alpha h}) \, dt. \qquad (45)$$

Proof. Since functions $u_{ht}, \lambda_{ht} \in L_\infty(Q_T)$ and functions $c_{\alpha h}, c_{glob} \in \widetilde{Y}$, then by (45) the function $D(c_{\alpha h}) \in L_\infty(\Omega)$. Next, using (41) and (45), we obtain the following approximate error estimate

$$
\begin{aligned}
|E(c_\alpha) - E(c_{h\alpha})| &\leq |E'(c_{\alpha h})(c_\alpha - c_\alpha^I)| \\
&\leq |D(c_{\alpha h})(c_\alpha - c_\alpha^I)| + |[E'(c_{\alpha h}) - D(c_{\alpha h})](c_\alpha - c_\alpha^I)| \\
&= |J_1| + |J_2|
\end{aligned} \qquad (46)
$$

It follows from (36), (45) and (46) that

$$|J_1| \leq \left(A(\Omega) K \|D(c_{\alpha h})\|_{L_\infty(\Omega)}\right)\|\nabla c_\alpha\|_{L_\infty(\Omega)} h. \qquad (47)$$

We now estimate $|J_2|$ in (46). It follows from (36), (39), (45) and (46) that

$$|J_2| \leq K\|\nabla c_\alpha\|_{L_\infty(\Omega)} h \int_\Omega \left|\int_0^T [(u_t\lambda_t) - (u_{ht}\lambda_{ht})](x, t, c_{\alpha h}) \, dt\right| dx$$

We have $(u_t\lambda_t) - (u_{ht}\lambda_{ht}) = (u_t - u_{ht})\lambda_t + u_{ht}(\lambda_t - \lambda_{ht})$. Next, by (24) and (43)

$$\|u_h\|_{H^1(Q_T)} \leq \|u_h - u\|_{H^1(Q_T)} + \|u\|_{H^1(Q_T)} \leq \zeta + B\|F\|_{H^5(Q_T)}.$$

Hence, since $\zeta \in (0,1)$, we obtain, using the Cauchy-Schwarz inequality that

$$\int_\Omega \left|\int_0^T [(u_t\lambda_t) - (u_{ht}\lambda_{ht})](x, t, c_{\alpha h}) \, dt\right| dx$$

$$\leq \|u - u_h\|_{H^1(Q_T)}\|\lambda\|_{H^1(Q_T)} + \|u_h\|_{H^1(Q_T)}\|\lambda - \lambda_h\|_{H^1(Q_T)} \leq C\zeta.$$

Hence,

$$|J_2| \leq CK\|\nabla c_\alpha\|_{L_\infty(\Omega)} \zeta h.$$

Combining this with (46) and (47), we obtain (44). $\square$

Remark 4.1. It follows from (43) that if solutions of state and adjoint problems are computed with a good accuracy, then the term $P(c_{\alpha h}) := A(\Omega) \|D(c_{\alpha h})\|_{L_\infty(\Omega)}$ dominates the term $C\zeta$ in (44), provided, of course that the number $P(c_{\alpha h})$ is not too small. The term $KP(c_{\alpha h}) \|\nabla c_\alpha\|_{L_\infty(\Omega)} h$ is also presented in a posteriori error estimates of [5, 6]. Although it was observed numerically in [5, 6] that this term dominates the rest of certain integral terms, this observation was not explained analytically.

Mesh Refinement Recommendation. Assume that conditions of Theorem 4.2 hold and that the function $D(c_{\alpha h})(x) \in C(\overline{\Omega})$. It follows from this theorem and Remark 4.1 that the mesh should be refined in such a subdomain of the domain Ω where values of the function $|D(c_{\alpha h})(x)|$ are close to the number

$$\max_{\overline{\Omega}} |D(c_{\alpha h})(x)| = \max_{\overline{\Omega}} \left| \alpha (c_{\alpha h} - c_{glob})(x) - \int_0^T (u_{ht}\lambda_{ht})(x, t, c_{\alpha h}) \, dt \right|.$$

5 The Adaptive Algorithm

In this section we outline our adaptive algorithm using the mesh refinement recommendation of section 4. Although this recommendation was actually derived from Theorems 4.1 and 4.2, which in turn assume that initial conditions (2) for equation (1) are replaced with initial conditions (12), we now present the adaptive algorithm for the case of generic initial conditions for equation (1). We refer to the beginning of section 3 for some discrepancies. In section 6 we use the adaptive algorithm of the current section. Our numerical experience of this and past publications [5, 6] shows that this is sufficient.

So, on each mesh we should find an approximate solution of the equation $D(c_{\alpha h}) = 0$. In other words, we should approximately solve the following equation with respect to the function $c_{\alpha h}(x)$,

$$\alpha (c_{\alpha h} - c_{glob})(x) - \int_0^T (u_{ht}\lambda_{ht})(x, t, c_{\alpha h}) \, dt = 0. \tag{48}$$

For each new mesh we first linearly interpolate the function $c_{glob}(x)$ on it. Since this function $c_{glob}(x)$ is computed on the globally convergent stage as a linear combination of finite elements of the initial mesh and these finite elements are piecewise linear functions, then subsequent linear interpolations on finer meshes do not change the function $c_{glob}(x)$. On each mesh we iteratively update approximations c_h^n of the function $c_{\alpha h}$. To do so, we use the quasi-Newton method with the classic BFGS update formula with the limited storage [26]. Denote

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

where functions $u_h(x, t, c_h^n), \lambda_h(x, t, c_h^n)$ are computed via solving state and adjoint problems with $c := c_h^n$.

Based on the mesh refinement recommendation of section 4, we use the following adaptivity algorithm in our computations:

Step 0. Choose an initial mesh K_h in Ω_1 and an initial time partition J_0 of the time interval $(0,T)$. Start with the initial approximation $c_h^0 = c_{glob}$ and compute the sequence of c_h^n via the following steps:

Step 1. Compute solutions $u_h(x, t, c_h^n)$ and $\lambda_h(x, t, c_h^n)$ of state and adjoint problems of (19) and (20) on K_h and J_k.

Step 2. Update the coefficient $c := c_h^{n+1}$ on K_h.

Step 3. Stop computing c_h^n if either $||g^n||_{L_2(\Omega_1)} \leq \theta_1$ or norms $||g^n||_{L_2(\Omega_1)}$ are stabilized. Otherwise set $n := n + 1$ and go to step 1. Here θ_1 is the tolerance in quasi-Newton updates. In our computations we took $\theta_1 = 10^{-5}$.

Step 4. Compute the function $B_h(x)$,

$$B_h(x) = \left| \alpha \left(c_{\alpha h} - c_{glob} \right)(x) - \int_0^T \left(u_{ht} \lambda_{ht} \right)(x, t, c_{\alpha h}) \, dt \right|.$$

Next, refine the mesh at all points where

$$B_h(x) \geq \beta_1 \max_{\overline{\Omega}} B_h(x). \tag{49}$$

Here the tolerance number $\beta_1 \in (0, 1)$ is chosen by the user, see section 6 for details.

Step 5. Construct a new mesh K_h in Ω_1 and a new time partition J_k of the time interval $(0, T)$. On J_k the new time step τ should be chosen in such a way that the CFL condition is satisfied. Interpolate the initial approximation c_{glob} from the previous mesh to the new mesh. Next, return to step 1 and perform all above steps on the new mesh.

Step 6. Stop mesh refinements if norms defined in step 3 either increase or stabilize, compared with the previous mesh, see Table 1 in section 6 for details.

6 Numerical Studies

6.1 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 with the given function $c(x)$. To solve the forward problem, we use the hybrid FEM/FDM method described in [7]. The computational domain for the forward problem in our test is $G = [-4.0, 4.0] \times [-5.0, 5.0]$. This domain is

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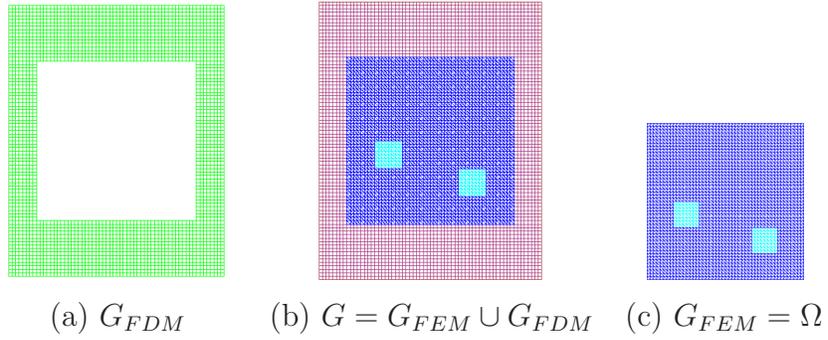
Figure 1: *The hybrid mesh (b) is a combinations of a structured mesh (a), where FDM is applied, and a mesh (c), where we use FEM, with a thin overlapping of structured elements. The solution of the inverse problem is computed in the square Ω and $c(x) = 1$ for $x \in G \setminus \Omega$.*

$t = 0.5$ $t = 3.7$ $t = 5.9$ $t = 6.9$

$t = 7.5$ $t = 8.5$ $t = 9.6$ $t = 11.2$
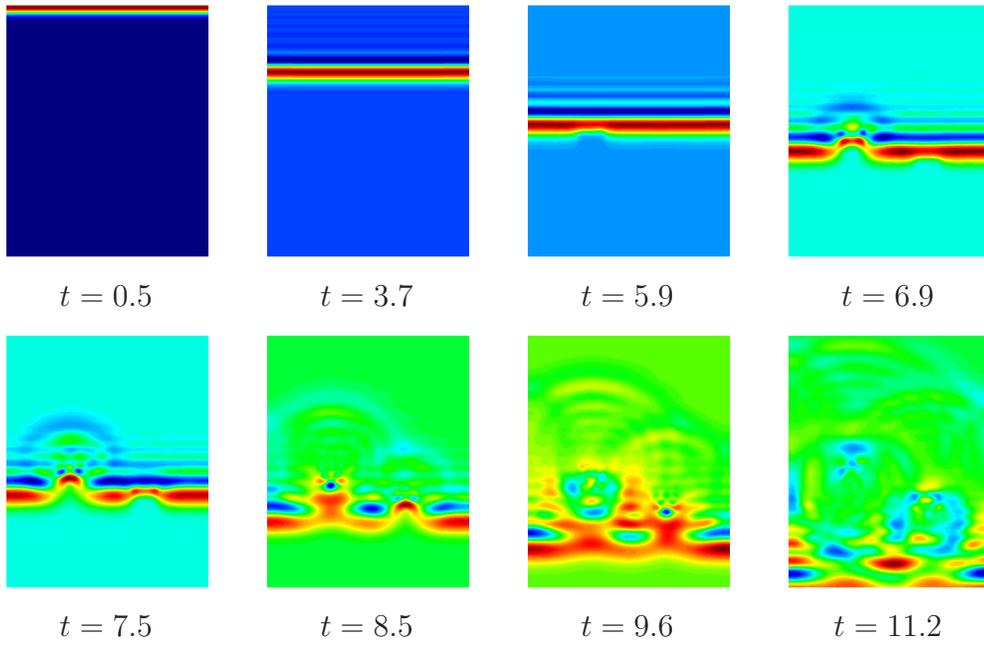
Figure 2: *Isosurfaces of the simulated exact solution to the forward problem (51) at different times with a plane wave initialized at the top boundary.*

split into a finite element domain $G_{FEM} := \Omega = [-3.0, 3.0] \times [-3.0, 3.0]$ and a surrounding domain G_{FDM} with a structured mesh, $G = G_{FEM} \cup G_{FDM}$, see Figure 1. The reason of the use of the hybrid method is that since we know that

$$c(x) = 1 \text{ in } G \setminus \Omega, \tag{50}$$

then there is no point to have a locally fine mesh in $G \setminus \Omega$. On the other hand, since inhomogeneities are located inside of Ω, then it is natural to have a locally fine mesh in Ω, which is provided by finite elements. The space mesh in Ω consists of triangles and in G_{FDM} - of squares with the mesh size $\widetilde{h} = 0.125$ in the overlapping regions. The boundary of the domain G is $\partial G = \partial G_1 \cup \partial G_2 \cup \partial G_3$. Here, ∂G_1 and ∂G_2 are respectively top and bottom sides of the largest domain of Figure 1 and ∂G_3 is the union of left and right sides of this domain. At ∂G_1 and ∂G_2 we use first-order absorbing boundary conditions [14]. At the lateral boundaries, mirror boundary conditions allow us to assume an infinite space domain in the lateral direction. 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 $g(x, t)$, see (5).

Thus, the forward problem in our test is

$$
\begin{aligned}
c(x) u_{tt} - \triangle u &= 0, \quad \text{in } G \times (0, T), \\
u(x, 0) &= 0, \ u_t(x, 0) = 0, \text{ in } G, \\
\partial_n u\big|_{\partial G_1} &= f(t), \text{ on } \partial G_1 \times (0, t_1], \\
\partial_n u\big|_{\partial G_1} &= -\partial_t u, \text{ on } \partial G_1 \times (t_1, T), \\
\partial_n u\big|_{\partial G_2} &= -\partial_t u, \text{ on } \partial G_2 \times (0, T), \\
\partial_n u\big|_{\partial G_3} &= 0, \text{ on } \partial G_3 \times (0, T),
\end{aligned}
\tag{51}
$$

where $f(t)$ is the plane wave defined as

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

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 [14] are used on top $\partial G_1 \times (t_1, T]$ and bottom $\partial G_2 \times (0, T]$ boundaries, and the Neumann boundary condition is used on ∂G_3. Figures 2 show how the plane wave propagates for the structure given on Fig. 1.

6.2 Reconstruction by the globally convergent algorithm

In our numerical experiment we reconstruct the medium, which is homogeneous with $c(x) = 1$ except of two small squares, where $c(x) = 4$, see Figure 1-c). However, we have not assumed *a priori* knowledge of neither the structure of this medium nor of the background constant $c(x) = 1$ outside of those two small squares. Although, following the Tikhonov concept (as mentioned in section 2), we have assumed the knowledge of the lower bound $c(x) \geq 1$ and

20

also that outside of the domain of interest Ω our function $c(x) = 1$, see (3) and (50). The latter is the reason why our starting value for the tail function is $V_{1,1}(x) = \overline{s}^{-2} \ln \widetilde{w}(x, \overline{s})$, where $\widetilde{w}(x, \overline{s})$ is the function $w(x, s)$ at $s = \overline{s}$ for the case $c \equiv 1$ (section 2). The assumption $c(x) \geq 1$ was used as follows: if at any point x' the reconstructed value of the coefficient was $c_{n,k}(x') \leq 1/2$, then we have used a new value at this point by setting $\widetilde{c}_{n,k}(x') := 1$. The latter has ensured that when computing the forward problem (51) with $c := c_{n,k}$, the corresponding operator is a hyperbolic one.

It was found in [4] that the interval $[\underline{s}, \overline{s}] = [6.7, 7.45]$ is the optimal one for domains G, Ω specified above, and so we have used it in our computations. We have chosen the step size with respect to the pseudo frequency $\rho = 0.05$. Hence, $N = 15$ in our case. We have chosen two sequences of regularization parameters $\mu := \mu_n$ and $\varkappa = \varkappa_n$ for $n = 1, ..., N$, which are the same as ones in [4]. So, values of these parameters as well as the value of the regularization parameter α in the adaptivity were:

$$\mu_n = 20, n = 1, 2; \mu_n = 200, n = 3, 4, 5; \mu_n = 2000, n \geq 6;$$
$$\varkappa_n = 0, n = 1, 2; \varkappa_n = 0.001, n = 3, 4, 5; \varkappa_n = 0.01, n = 6, 7; \varkappa_n = 0.1 \text{ for } n \geq 8; \alpha = 0.01.$$

Once the function $q_{n,k}$ is calculated, we update the function $c := c_{n,k}$ via backwards calculation, see subsection 7.3 of [4] for some numerical details. The resulting computed function is $c_{glob}(x) := c_{\overline{N}}(x)$. In the current work we choose a stopping rule which is completely different from one in [4]. We have observed that the lower boundary Γ of the square Ω,

$$\Gamma = \{x_2 = -3\} \cap \overline{\Omega}. \tag{52}$$

is such a part of the boundary $\partial\Omega$, which is the most sensitive one to the presence of inclusions. So, denote $\Gamma_{\widetilde{h}} = \left\{ (x_1, x_2) \in \Omega : x_2 = -3 + \widetilde{h} \right\}$. In other words, $\Gamma_{\widetilde{h}}$ is the part of the horizontal straight line, which is inside of the square Ω, and the distance between $\Gamma_{\widetilde{h}}$ and the lower boundary $\{x_2 = -3\}$ of Ω is $\widetilde{h} = 0.125$. When calculating iterations with respect to the nonlinear term (section 2), we consider norms F_n^k,

$$F_n^k = ||q_n^k|_{\Gamma_{\widetilde{h}}} - \overline{\psi}_n||_{L_2(-3,3)}.$$

We stop our iterations with respect to the nonlinear term when

$$\text{either } F_n^k \geq F_n^{k-1} \text{ or } F_n^k \leq \upsilon, \tag{53}$$

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

$$\text{either } F_{n,k} \geq F_{n,k-1} \text{ or } F_{n,k} \leq \upsilon, \tag{54}$$

where $F_{n,k} = ||q_{n,k}|_{\Gamma_{\widetilde{h}}} - \overline{\psi}_n||_{L_2(-3,3)}$. Recall that the number k, on which these iterations are stopped, is denoted as $k := m_n$. Once the convergence criterion (54) is satisfied, we take the

21

a) $c_{9,2}$ b) $c_{10,2}$ c) $c_{11,2}$
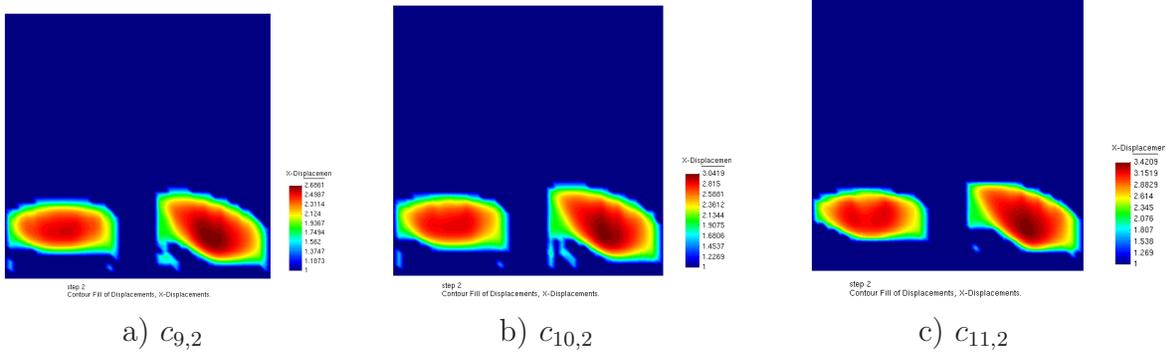
Figure 3: *Results of the performance of the globally convergent stage of our two-stage numerical procedure. Spatial distributions of some functions $c_{n,k}$. The function $c_{11,2}$ is taken as the final result of this stage (see details in the text). The maximal value of $c_{11,2}(x) = 3.8$ within each imaged inclusion. Also, $c_{11,2}(x) = 1$ outside of both imaged inclusions. Hence, the $3.8 : 1$ inclusion/background contrast is imaged with only 5% of error (the correct one is $4 : 1$). However, while the location of the right inclusion is imaged accurately, comparison with Fig.1-c shows that the left imaged inclusion is located below its correct position, and so it is desirable to shift it upwards. This paves the way for a subsequent application of the adaptivity technique, which takes the function $c_{11,2} := c_{glob}$ as the starting point for computations.*

last computed tail V_{n,m_n}, set $V_{n+1,1} := V_{n,m_n}$ and run computations again for q_{n+1}, see section 2. Hence, the number m_n of iterations with respect to tails is chosen automatically "inside" of each iteration for q_n, which means that m_n varies with n. So, new convergence criteria (53) and (54) represent a more flexible stopping rule in the globally convergent algorithm compared with the one of [4], since numbers m_n where not chosen automatically in [4].

In our numerical test we have considered the noisy boundary data g_σ introduced as

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

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, ς_j is a random number in the interval $[-1; 1]$, g_{max} and g_{min} are respectively maximal and minimal values of the computed boundary data g in (5) 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).

Figure 3 displays results of the performance of the globally convergent stage of our two-stage numerical procedure. One can see that the location of the right small square is imaged well. It follows from Figure 3-c) that the imaged contrast in this square is $3.8 : 1 = \max c_{11,2} : 1$, where $n := \overline{N} = 11$ is our final iteration number (see below for this choice of $\overline{N}$). Thus, we have obtained the 5% error $(0.2/4)$ in the imaged contrast, which is the same as the error

a)   b)
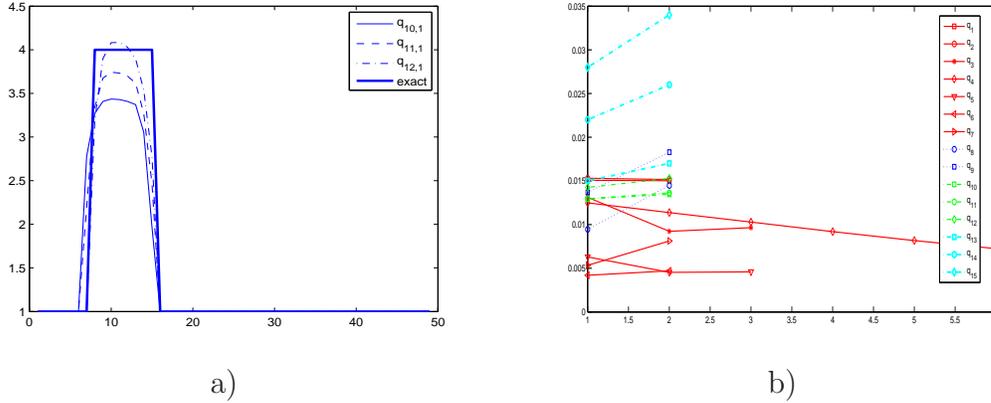
Figure 4: *a) The one-dimensional cross-sections of the image of the function $c_{n,k}$ computed for corresponding functions $q_{n,1}$ along the vertical line passing through the middle of the right small square; b) Computed L_2 -norms of the $F_{n,k} = ||q_{n,k}|_{\partial\Omega} - \overline{\psi}_n||_{L_2(-3,3)}$.*

in the input data. As to the left small square, we got the same $3.8 : 1$ contrast in it for $c_{11,2}(x)$. Values of the function $c_{11,2}(x) = 1$ outside of these squares are imaged accurately. However, the location of the left square is shifted downwards. So that both imaged squares are on about the same horizontal level. Therefore, comparison with Fig. 1-c) reveals that it is desirable to shift the left imaged square upwards. This opens the door for the subsequent application of the adaptivity technique.

Figure 4-b) shows computed L_2-norms $F_{n,k}$. Using this figure, we analyze results of our reconstruction. One can see on Figure 4-b) 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 (54). We observe that the norms $F_{n,k}$ generally decrease until computing the function q_7. Next, they slightly grow, decay from $F_{9,2}$ to $F_{10,2}$ and then these norms stabilize on $n = 11, 12$. For $n = 13, 14, 15$ these norms grow steeply. Thus, we conclude, that we should stop our iterations at $\overline{N} = 11$. So, we take the function $c_{11,2} := c_{glob}$ as our final reconstruction result on the globally convergent stage. We again refer to page 157 of [13] for an explanation of the idea of the choice of the iteration number as a regularization parameter.

6.3 Synthesis of the globally convergent algorithm with the adaptivity

In this subsection we demonstrate the performance of the synthesis of our globally convergent algorithm with the adaptivity technique. In our case the domain Ω_1 (section 3) is $\Omega_1 = \{x : x_2 > -3\} \cap G$. Hence, $\Omega \subset \Omega_1$ and the line Γ in (52) is a part of the lower boundary of the rectangle Ω_1. The normal derivative $\partial_n u |_{\partial\Omega_1}$, which we need for the solution of the state problem (19), was found via the solution of the forward problem (51) in the data simulation. We have not used the function $z_\varepsilon(t)$ in (20) because we have observed in our computations that $(\widetilde{g} - u)(x, t)|_{\partial\Omega_1} \approx 0$ for $t \approx T$.

23

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

Table 1: *Norms $\|u\mid_{\Gamma_T} -g\|_{L_2(\Gamma_T)}$ on adaptively refined meshes. Here $\Gamma_T = \Gamma \times (0,T)$ and n is the number of updates in the quasi-Newton method. These norms generally decrease as meshes are refined. Then they slightly increase on the 4th refinement. Thus, using this table, we conclude that on the four times refined mesh we get the final solution of our inverse problem, which corresponds to Figure 5-j.*

We have applied the adaptive algorithm of section 5. Now the question is on how to choose the tolerance number β_1 in (49). If we choose β_1 too small (for example, if $\beta_1 = 0$), then we will refine mesh in almost the entire domain Ω, since, realistically, the function $B_h(x)$ obtained in the optimization procedure is non- zero at almost all mesh points. Unlike this, our goal is to construct a new mesh with a few nodes as possible, while still getting a good refinement of the solution obtained on the globally convergent stage of our two-stage numerical procedure. On the other hand, the parameter β_1 cannot be taken too close to 1 either, since then the automatic adaptive algorithm would come up with a too narrow region, where the mesh should be refined. Thus, the choice of β_1 depends on concrete values of the function $B_h(x)$ and should be figured out in numerical experiments. So, we take $\beta_1 = 0.1$ on the initial coarse mesh, $\beta_1 = 0.2$ on the one, two and three times refined meshes, and $\beta_1 = 0.6$ for all follow up refinements of the initial mesh.

On all refined meshes we have used a cut-off parameter C_{cut} for the reconstructed coefficient $c_{\alpha h}$. So that we re-define $c_{\alpha h}$ as

$$c_{\alpha h}(x) := \begin{cases} c_{\alpha h}(x), & \text{if } |c_{\alpha h}(x) - c_{glob}(x)| \geq C_{cut} \\ c_{glob}(x), & \text{elsewhere.} \end{cases}$$

We choose $C_{cut} = 0$ for $m < 3$ and $C_{cut} = 0.3$ for $m \geq 3$, where m is the number of iterations in the quasi-Newton method on each mesh. Hence, the cut-off parameter ensures that we do not go too far from our good first guess for the solution $c_{glob}(x)$.

In the adaptive algorithm we can use box constrains for the reconstructed coefficient. We obtain these constraints using the solution computed on the globally convergent stage of our two-stage numerical procedure. Namely, minimal and maximal values of the target coefficient in box constraints are taken using results of the first stage. Since the function c_{glob} obtained on the first stage is a good approximation for the correct solution, and $c_{glob}(x) \in [1, 3.8]$, we enforce that the coefficient $c(x)$ belongs to the following set of admissible parameters $c \in P = \{c \in C(\overline{\Omega}) | 1 \leq c(x) \leq 4\}$.

First, the function $c_{glob}(x)$ was taken on the initial coarse mesh shown on Figure 5-a) and the quasi-Newton method was applied on this mesh. Comparison of Figures 5-d and 3-c (for

a) 4608 elements   b) 5340 elements   c) 6356 elements

d) 4608 elements   e) 5340 elements   f) 6356 elements

g) 10058 elements   h) 14586 elements

i) 10058 elements   j) 14586 elements
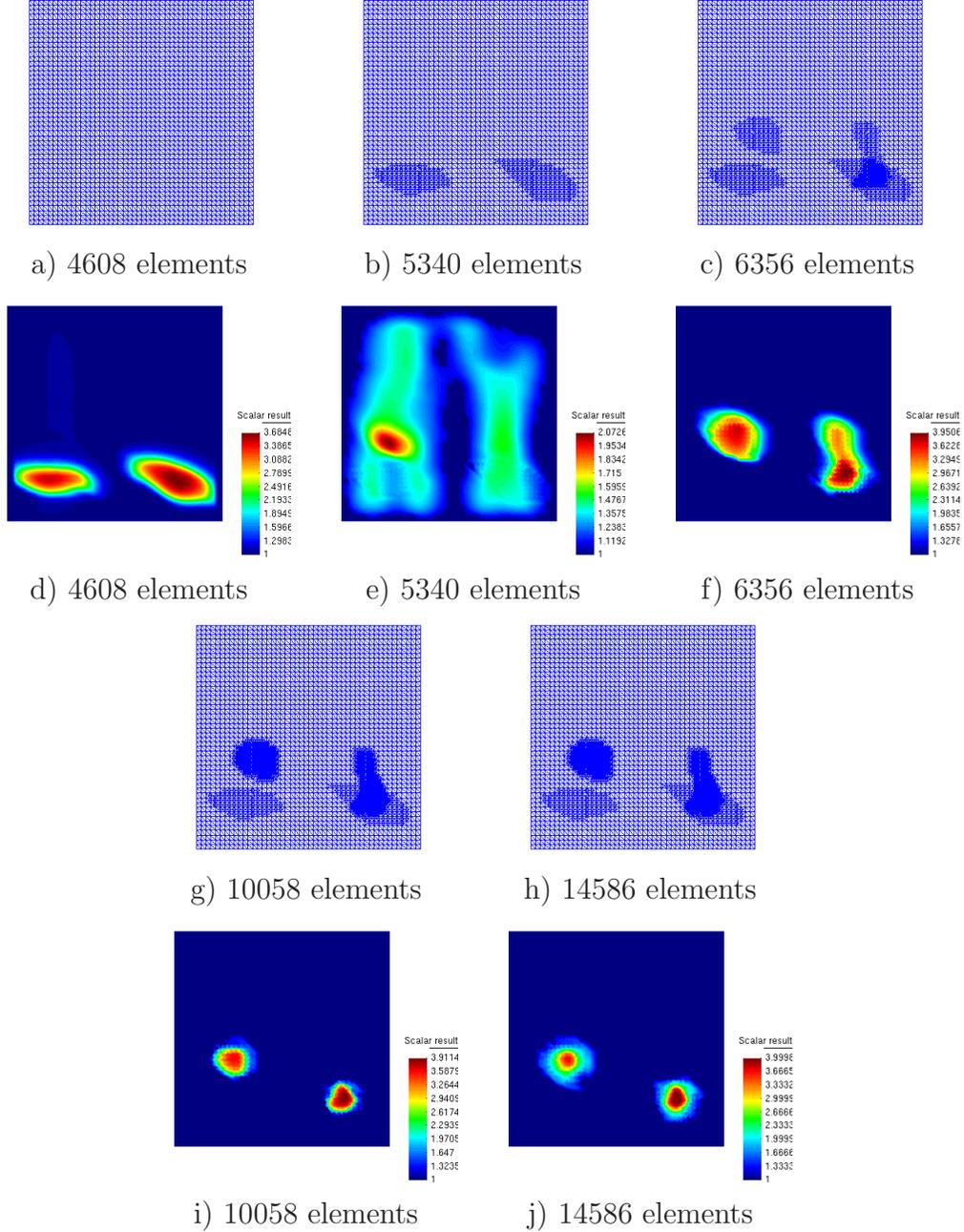
Figure 5: *Adaptively refined meshes a)-c),g),h) and corresponding images d)-f), i),j) on the second stage of our two-stage numerical procedure. In a) the same mesh was used as one on the globally convergent stage. Comparison of d) with Fig. 3-c) (for $c_{11,2} = c_{glob}$) shows that the image was not improved compared with the globally convergent stage when the same mesh was used. However, the image was improved due to further mesh refinements. Fig. j) displays the final image obtained after four mesh refinements. Locations of both inclusions as well as 4:1 inclusions/background contrasts in them are imaged accurately, see details in the text and compare with Fig. 1-c.*

$c_{11,2} := c_{glob}$) shows that the image was not improved, compared with the one obtained on the globally convergent stage. Next, the mesh was adaptively refined four times, using the procedure described in section 5. Adaptively refined meshes are shown on Figures 5-b,c), g) and h). Table 1 presents computed L_2-norms of $||u|_{\Gamma_T} - g||_{L_2(\Gamma_T)}$, where $\Gamma_T = \Gamma \times (0, T)$ and Γ was defined in (52). We observe that norms at the boundary generally decrease as meshes are refined. Then they slightly increase at the fourth refinement. Thus, using this table, we conclude that on the four times refined mesh we get the final solution of our inverse problem, which corresponds to Figure 5-j). One can see on Figure 5-j) that we are able to accurately reconstruct locations of both small squares. At the same time, an accurate inclusion/background contrast obtained on the globally convergent stage is preserved and even improved. This contrast is now 4:1=$\max c_f(x) : 1$ instead of 3.8:1 calculated on the first stage, where $c_f(x)$ is the final imaged coefficient. The value of the coefficient $c_f(x) = 1$ outside of small squares is also imaged well.

References

[1] Ainsworth M and Oden J T 2000 *A Posteriori Error Estimation in Finite Element Analysis* (New York: Wiley)

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

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

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

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

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

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

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

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

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

[11] Chen Y 1997 Inverse scattering and Heisenberg uncertainty principle *Inverse Problems* **13** 253-282

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

[13] Engl H W, Hanke M and Neubauer A 2000 *Regularization of Inverse Problems* (Boston: Kluwer Academic Publishers)

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

[15] Eriksson K, Estep D and Johnson C 2004 *Applied Mathematics: Body and Soul. Calculus in Several Dimensions* (Berlin: Springer)

[16] Evans L C 1998 *Partial Differential Equations, AMS, Providence* (Providence, RI : AMS)

[17] Griesbaum A, Kaltenbacher B, Vexler B 2008 Efficient computation of the Tikhonov regularization parameter by goal-oriented adaptive discretization *Inverse Problems* **24** 025025

[18] Kabanikhin S, Satybaev A and Shishlenin M 2004 *Direct Methods of Solving Multidimensional Inverse Hyperbolic Problems* (Utrecht, The Netherlands: VSP)

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

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

[21] Ladyzhenskaya O A and Uralceva N N 1969 *Linear and Quasilinear Elliptic Equations* (New York: Academic Press)

[22] Ladyzhenskaya O A 1985 *Boundary Value Problems of Mathematical Physics* (Berlin: Springer)

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

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

[25] Novikov R G 2005 The $\overline{\partial}$ approach to approximate inverse scattering at fixed energy in three dimensions *International Math. Research Papers* **6** 287-349

[26] Nocedal J 1991 Updating quasi-Newton matrices with limited storage, *Math. Comp.* **35** 773–782

[27] Tikhonov A N and Arsenin V Ya 1977 *Solutions of Ill-Posed Problems* (Washington, DC: Winston and Sons)