A globally convergent numerical method and the adaptivity technique for a hyperbolic coefficient inverse problem. Part I: analytical study.

Larisa Beilina

Department of Mathematical Sciences, Chalmers University of Technology and Gothenburg University, Gothenburg, SE-421196, Sweden

E-mail: `larisa.beilina@chalmers.se`

Michael V. Klibanov

Department of Mathematics and Statistics, University of North Carolina at Charlotte, Charlotte, NC 28223, USA

E-mail: `mklibanv@uncc.edu`

Abstract. A globally convergent numerical method for a multidimensional Coefficient Inverse Problem for a hyperbolic equation is presented. It is shown that this technique provides a good starting point for the so-called finite element adaptive method (adaptivity). This leads to a natural two-stage numerical procedure, which synthesizes both these methods. A new method for obtaining a posteriori error estimates for the adaptivity tecxhnique is demonstrated on a specific example of a hyperbolic Coefficient Inverse Problem.

1. Introduction

This paper is a continuation of the previous publication of the authors [5], where a new globally convergent numerical method for a Coefficient Inverse Problem (CIP) for a hyperbolic PDE was developed. In this first part of our work analytical studies are presented and the second part [8] discusses numerical experiments. Since the globally convergent numerical method was described in [5], we focus our analytical effort here on the adaptivity technique. We present a new idea of obtaining a posteriori error estimates both for the Lagrangian and for the regularized coefficient. Although this idea can probably be presented in a rather abstract form, we intentionally do not do this here preferring its demonstration for a specific CIP. Still, we outline a more general framework in the end of this first part. It is likely that that this framework might be extended to the parameter identification problem.

The CIP of this publication can be applied to inverse scattering of acoustical and electromagnetic waves. Compared with [5], the main new element here is a synthesis of the method of [5] with the locally convergent so-called Finite Element Adaptive method, which we call "adaptivity" for brevity. The adaptivity technique for inverse problems was previously developed in [4, 9, 10, 11]. The underlying reason of this synthesis is that the estimate of the difference between the correct solution and the computed one in the global convergence theorem of [5] depends on a small positive parameter η. This parameter incorporates both the error in the boundary data and errors generated by some approximations of the numerical procedure of [5]. The error in the boundary data models the error in measurements and is, therefore, unavoidable. At the same time, two other approximation errors cannot be made zero and they are not parts of previously developed locally convergent techniques. On the other hand, since η is small, then the global convergence theorem [5] guarantees that the solution obtained by the technique of [5] provides a good approximation for the correct solution of the CIP. Therefore, this solution can be used as a good starting point for a subsequent enhancement via a locally convergent numerical method. It was shown in previous publications [4, 9, 10, 11] that the adaptivity is capable to provide good quality images if a good first approximation for the solution is available. The latter leads to a logical conclusion that a synthesis of the adaptivity with the globally convergent method of [5] should be used. 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 addition to this two-stage procedure, there are five (5) new elements of this paper compared with [5]. We now list three of them. Two others are related to the adaptivity and are outlined below in this section. (1) The globally convergent algorithm is different from one in [5] in the sense that now "inner" iterations with respect to terms in certain quasilinear elliptic equations are used until they converge. Whereas previously a priori chosen number of iterations was used. (2) The stopping rule for the globally convergent part differs from one of [5]. Namely, we now evaluate certain L_2 norms at the boundary rather than inside of the domain of interest. (3) 2-D numerical examples are different from ones of [5].

We call a numerical method for a CIP *globally convergent* if: (1) a theorem is proven, which ensures that this method leads to a good approximation for the correct solution of that CIP, regardless on the availability of a priori given good first guess for that solution, and (2) this theorem is confirmed by numerical experiments. On the other hand, convergence of a locally convergent numerical method to the correct solution can be guaranteed only if the starting point is located in a small neighbourhood of this solution. The method of [5] relies on the structure of the underlying PDE operator instead of conventional least squares minimization techniques. This helps to avoid the phenomenon of local minima.

The adaptivity minimizes least squares objective functionals on a sequence of adaptively refined meshes until images are stabilized. On each mesh the minimization is performed via the quasi-Newton method. Convergence of Newton-like methods for general ill-posed problems was proven in [3]. At the same time, our numerical experiments demonstrate that just a straightforward application of the quasi-Newton method, which works on the same mesh as the globally convergent part did, does not improve the result obtained on the globally convergent stage. On the other hand, further adaptive mesh refinements indeed enhance the solution. Therefore it is *important* to use the adaptivity in our two-stage numerical procedure. In this paper we present a new analytical framework for the adaptivity, which is a modification of the framework of previous publications [9, 10, 11].

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 forward problems these upper bounds are obtained from classic a priori estimates of solutions. In the case of CIPs upper estimates are assumed to be known in advance, which is according to the Tikhonov concept for ill-posed problems [23, 25].

A posteriori error analysis addresses the main question of the adaptivity: *where to refine the mesh?*. This analysis provides upper estimates for differences between computed and exact solutions locally, in subdomains of the original domain. Such an analysis is a classic tool in applications of the adaptivity to forward problems for PDEs, see, e.g. [1, 16]. 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 changes the situation *radically*. In fact, the ill-posedness represents the major obstacle for an estimate of the difference between computed and exact coefficients. For this reason, the accuracy of the Lagrangian (depending on that coefficient) is usually estimated instead of one of the target coefficient [4, 9, 10, 11].

In this paper we develop a new idea for the derivation of a new a posteriori error estimate for the Lagrangian. This estimate is both stronger and more effective than one in [9, 10, 11]. The meaning of this estimate is that it indicates that one can ignore certain integral terms in the Frechet derivative of the Lagrangian when deciding where to refine the mesh. While these integral terms were also ignored in numerical experiments in [9, 10, 11] due to computational observations, an analytical explanation was not provided in these references. At the same time, it is more desirable to obtain a posteriori error estimate for the target coefficient rather than for the Lagrangian only. In [11] a heuristic estimate of such sort was obtained under the assumption of the existence of a solution of a certain problem for the Hessian. This existence was demonstrated numerically in [11].

In this paper we obtain the above mentioned a posteriori error estimate rigorously. We now specify this statement. The Lagrange functional is a modified classic Tikhonov regularization functional, see, e.g. [14, 23, 25] for the latter. Hence, similarly with [14, 23, 25], we call a minimizer of the Lagrangian for our CIP as the "regularized coefficient". Thus, given a value of the regularization parameter, we estimate the difference between the regularized coefficient and its approximation obtained on a certain finite element mesh. We assume in our derivation the existence of the solution of a certain problem for the Hessian. The proof of a corresponding existence theorem is a quite challenging problem, which is outside of the scope of this publication. The resulting a posteriori estimate differs from one of the Lagrangian only by a constant multiplier. This means that via refining mesh "for the Lagrangian", we actually obtain a better accuracy for the regularized coefficient.

It would be better of course to provide a posteriori estimate of the difference between computed and exact (rather than regularized) coefficients. It is well known, however that even an estimate of the distance between regularized and exact solutions of a CIP is a very challenging and still unsolved (in many cases) problem. Indeed, this problem requires a derivation of an upper estimate of the modulus of the continuity, on a compact set, of the operator, which is inverse to the operator of the CIP, see (2.6) in §1 of Chapter 2 of [23]. Thus, the problem of an estimate of the distance between computed and exact solutions is not considered here. Still, it follows from Theorem 2 on p. 65 of [25] that one can often guarantee that the regularized solution is close to the exact one, although without an explicit estimate. Hence, assuming that this is the case, our a posteriori error estimate can be considered as an approximate estimate of the distance between computed and exact coefficients.

The first part of our work is organized as follows. In section 2 we briefly describe the globally convergent method, see [5] for details. In section 3 we present a modified framework for the adaptivity technique and prove a posteriori error estimates. A preprint with this publication is available online [7]. One can also find there a preliminary preprint [6] where more numerical results are available.

2. Brief description of the globally convergent numerical method of [5]

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

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

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

Since equation (1) governs a wide range of applications, including e.g., propagation of acoustic and electromagnetic waves, then the same is true for the CIP we consider. In the acoustical case $1/\sqrt{c(x)}$ is the sound speed. In the 2-D case of EM waves propagation in a non-magnetic medium, the dimensionless coefficient $c(x) = \varepsilon_r(x)$, where $\varepsilon_r(x)$ is the relative dielectric function of the medium, see [12], where this equation was derived from Maxwell's equations in the 2-D case. Let d_1 and d_2 be two positive constants and $\Omega \subset \mathbb{R}^3$ be a convex bounded domain with the boundary $\partial\Omega \in C^3$. We assume that the coefficient $c(x)$ of equation (1) is such that

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

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

We consider the following

Inverse Problem. Suppose that the coefficient $c(x)$ satisfies (3) and (4), where the positive numbers d_1 and d_2 are given. Assume that the function $c(x)$ is unknown in the domain Ω. Determine the function $c(x)$ for $x \in \Omega$, assuming that the following function $g(x,t)$ is known for a single source position $x_0 \notin \overline{\Omega}$

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

A priori knowledge of constants d_1, d_2 corresponds well with the Tikhonov concept about the availability of a priori information for an ill-posed problem [25]. In applications the assumption $c(x) = 2d_1$ for $x \in \mathbb{R}^3 \diagdown \Omega$ means that the target coefficient $c(x)$ has a known constant value outside of the medium of interest Ω. Another argument here is that one should bound the coefficient $c(x)$ from the below by a positive number to ensure that the operator in (1) is a hyperbolic one on all iterations of our method. Since we do not impose any "smallness" conditions on numbers d_1 and d_2, our numerical method is not a locally convergent one. The function $g(x,t)$ models time dependent measurements of the wave field at the boundary of the domain of interest. In practice measurements are performed at a number of detectors, of course. In this case the function $g(x,t)$ can be obtained via one of standard interpolation procedures, a discussion of which is outside the scope of this publication. In the case of a finite time interval, on which measurements are performed, one should assume that this interval is large enough and thus, the t-integral of the Laplace transform over this interval is approximately the same as one over $(0,\infty)$.

The question of uniqueness of this Inverse Problem is a well known long standing open problem. It is addressed positively only if the function $\delta(x - x_0)$ above is replaced with a such a function $f(x) \in C^\infty(\mathbb{R}^3)$ that $f(x) \neq 0, \forall x \in \overline{\Omega}$. An example of this function is

$$f_\varepsilon(x) = C_\varepsilon e^{-\frac{|x-x_0|^2}{\varepsilon^2}}, \quad \int_{\mathbb{R}^3} f_\varepsilon(x)\,dx = 1, \tag{6}$$

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

Consider the Laplace transform of the functions u,

$$w(x,s) = \int_0^\infty u(x,t)e^{-st}dt, \text{ for } s > \underline{s} = const. > 0, \tag{7}$$

where $\underline{s}$ is a certain number. It is sufficient to choose $\underline{s}$ such that the integral (7) 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. Since by the maximum principle $w(x,s) > 0$, then we can consider the function $q(x,s) = \partial_s(s^{-2} lnw(x,s))$. This function satisfies a certain nonlinear integral differential equation with Volterra integrals with respect to s, where integration is carried out from s to $\bar{s}$, where $\bar{s}$ is the value of the pseudo frequency at which these integrals are truncated. In that equation the so- called tail function is also involved. This function complements that truncation, it is unknown and it is small because of a certain asymptotic behaviour at $\bar{s} \to \infty$. Therefore that equation contains two unknown functions q and the tail. The reason why we can approximate both of them is that we treat them differently: while the function q is approximated via inner iterations, the tail function is approximated via outer iterations. Consider a partition of the interval into small subintervals with the length of h. Approximate the function q as a piecewise constant function q_n with respect to s on each of these small intervals $(s_n, s_{n-1}]$. Next, the equation for q_n is multiplied by the Carleman Weight Function $CWF = e^{\mu(s-s_{n-1})}$, where μ is a large parameter. Then the resulting equation is integrated with respect to $s \in (s_n, s_{n-1}]$. As a result, a finite sequence of Dirichlet boundary value problems for nonlinear elliptic PDEs for functions q_n is obtained, where Dirichlet boundary conditions are known. This system is solved sequentially. As soon as the function q_n is approximated, an approximation c_n for the unknown coefficient c is found and the next update for the tail function is also found. The first approximation for the tail is either zero or the one which corresponds to the solution of the above Cauchy problem for $c = 2d_1$. Let σ be a small parameter characterizing the level of the error in the data, and ϵ be a certain small regularization parameter which is introduced to improve the stability property of solving the above Dirichlet boundary value problems. Let $\xi > 0$ be a small number such that certain norm of the tail is less than ξ. Denote $\eta = 2(h + \sigma + \epsilon + \xi)$. Then η is small. The global convergence theorem of [5] claims that $|c_n - c^*|_\alpha \leq C\eta$, where $|\cdot|_\alpha$ is a Hölder norm, c^* is the exact solution of our CIP satisfying (1), (2) and $C > 0$ is a constant. Thus, the globally convergent part provides a good approximation for the exact solution.

3. The Adaptivity

3.1. State and adjoint problems

To use the adaptivity technique, we formulate our CIP inverse problem as an optimization problem, where we seek the unknown coefficient $c(x)$, which gives the solution of the boundary value problem (1), (2) for the function $u(x,t)$ with the best least squares fit to the time domain observations $g\left(x,t\right)$, see (5). Denote $Q_T = \Omega \times (0,T), S_T = \partial\Omega \times (0,T)$. In this section $C = C\left(Q_T\right)$ denotes different positive constants depending only on the domain Q_T, in (15) C depends only on Ω. Our goal now is to find the function $c(x)$ which minimizes the Tikhonov functional

$$E(u,c) = \frac{1}{2}\int\limits_{S_T} (u\mid_{S_T} - g(x,t))^2 \zeta_{\varepsilon_1}(t)\, d\sigma dt + \frac{1}{2}\gamma \int\limits_{\Omega} (c - c_0)^2 \ \ dx, \tag{8}$$

where $\gamma \in (0,1)$ is a small regularization parameter, c_0 is an initial guess for the unknown coefficient c and the function $\zeta_{\varepsilon_1}(t)$ is introduced in order to ensure compatibility conditions in the so-called "adjoint problem" (below). This function has the following properties at $t = T$

$$\zeta_{\varepsilon_1} \in C^3\left[0,T\right], \zeta_{\varepsilon_1}(t) = \left\{ \begin{array}{c} 0, t \in (T - \varepsilon_1/2, T) \\ 1, t \in (0, T - \varepsilon_1), \\ \text{between 0 and } 1, t \in (T - \varepsilon_1, T - \varepsilon_1/2) \end{array} \right\}.$$

Here $\varepsilon_1 \in (0,T)$ is a sufficiently small number. The existence of such functions is known from The Real Analysis course. In principle, the regularization theory requires that the norm of the penalty term in (8) should be stronger than the $L_2(\Omega)$ norm [14, 23, 25]. However, the stronger norm condition is sufficient but not necessary. Thus, we use the simpler $L_2(\Omega)$ norm here, because our computational experience shows that this is sufficient for our CIP. Another justification of this is that all norms are equivalent in finite dimensional spaces, with which we actually work in our computations.

Since $c(x) = 1$ in $\mathbb{R}^3 \diagdown \Omega$, then given the function $g(x,t)$ in (5), one can uniquely determine the function $u(x,t)$ for $(x,t) \in (\mathbb{R}^3 \diagdown \Omega) \times (0,T)$ as the solution of the boundary value problem for equation (1) with initial conditions (2) and with the boundary condition (5). Hence, one can uniquely determine the function $p(x,t) = \partial_n u\mid_{S_T}$. Thus, in this section we consider initial boundary value problems only in the domain Q_T. In particular, the function u of (1), (2) is the solution of the following so-called "state problem"

$$\begin{aligned} cu_{tt} - \triangle u &= 0, (x,t) \in Q_T, \\ u(x,0) &= u_t(x,0) = 0, \\ \partial_n u\mid_{S_T} &= p(x,t). \end{aligned} \tag{9}$$

In this section we consider the following condition imposed on the function $c(x)$

$$c \in C\left(\overline{\Omega}\right) \cap H^1(\Omega), c_{x_i} \in L_\infty(\Omega), i = 1,2,3; \quad c(x) \in [d_1, 2d_2] \text{ in } \Omega. \tag{10}$$

Note that first two relations of (10) are always in place if the function $c(x)$ is represented via a linear combination of standard piecewise linear finite elements. In addition, in order to guarantee that solutions of state and adjoint problems belongs to $H^2(Q_T)$, we assume that there exist such functions P, G and cut-off function a that

$$P, G \in H^3(Q_T), \partial_n P\mid_{S_T} = p(x,t), P(x,0) = P_t(x,0) = 0, \partial_n G\mid_{S_T} = g(x,t), \tag{11}$$

$$a \in C^\infty\left(\overline{\Omega}\right), a\mid_{\partial\Omega} = 1, \partial_n a\mid_{\partial\Omega} = 0. \tag{12}$$

For example, if $\Omega = \{|x| < R\}$, then $a(x) = \left[(|x| - R)^2 + 1\right]\chi(|x|)$, where the function χ is such that $\chi(z) \in C^\infty[0,R], \chi(z) = 1$ for $z \in [R/2, R], \chi(z) = 0$ for $z \in [0, R/4]$, and $\chi \in [0,1]$ for $z \in [R/4, R/2]$. Although functions $\sigma(x)$ might also be constructed for a more general domain, we are not doing this here for brevity. The existence of functions P, G satisfying (11) cannot be guaranteed for the case of the fundamental solution (1), (2). To guarantee (11), one could replace, for example the $\delta(x - x_0)$ function in (2) with a function $\widetilde{f_\varepsilon}(x) \in C^\infty(\mathbb{R}^3)$, similar with the function $f_\varepsilon(x)$ in (6) with the only difference that $\widetilde{f_\varepsilon}(x) = 0$ in Ω. Then the existence of functions P, G satisfying (11) can be guaranteed at least for the case when functions p, g are given without a random error. The question of an extension of the above globally convergent numerical method on the case of such replacement of the initial condition is outside of the scope of the current publication. Overall, the question of the existence of functions P, G satisfying (11) is one of discrepancies between our theory and computational practice, see part II [8] for more discrepancies.

Using (10), conditions (11) for P and applying slight modifications of proofs of either Theorems 4.1 and 5.1 of Chapter 4 of [22] or of Theorem 5 of section 7.2 of [17], we obtain that there exists unique solution $u \in H^2(Q_T)$ of the problem (9). Furthermore,

$$\partial_t^k u \in L_\infty\left(0, T; H^{2-k}(\Omega)\right), k = 0, 1, 2, \tag{13}$$

where $H^0(\Omega) := L_2(\Omega)$. In addition, the following integral identity holds

$$\int\limits_{Q_T} \left(-c(x)\, u_t r_t + \nabla u \nabla r\right) dxdt - \int\limits_{S_T} p r d\sigma dt = 0, \forall r \in H^1(Q_T), r(x, T) = 0. \tag{14}$$

We note that (14) is also the definition of the weak $H^1(Q_T)-$solution of the problem (9). The existence and uniqueness of this solution is guaranteed if the function P satisfies a weaker smoothness condition $P \in H^2(Q_T)$, see Theorem 5.1 of Chapter 4 of [22].

Denote

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

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 formulate the FEM for boundary value problems below, we introduce finite element spaces $W_h^u \subset H_u^1(Q_T)$ and $W_h^\varphi \subset H_\varphi^1(Q_T)$. These spaces consist of standard piecewise linear finite elements 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 space $V_h \subset L_2(\Omega)$ of standard piecewise linear finite elements for the target coefficient $c(x)$ and denote $U_h = W_h^u \times W_h^\varphi \times V_h$. Obviously $U_h \subset \bar{U}$ as a set. So, we consider U_h as a discrete analogue of the space $\bar{U}$. It is convenient for us to introduce in U_h the same norm as one in $\bar{U}^1$, $\|\bullet\|_{U_h} := \|\bullet\|_{\bar{U}^1}$. We work with piecewise linear finite elements in our analytical derivations because we work with them in numerical experiments. Considerations of other types of finite elements are outside of the scope of this publication. We assume below that the mesh in the domain Ω is regular.

We now formulate some error estimates for interpolants in the format, which is convenient for our derivations below. Let h and τ be maximal grid step sizes of piecewise linear finite elements with respect to x and t respectively. For any function f belonging to either $H^2(Q_T)$ or to $H^1(\Omega)$, let f^I be its interpolant via corresponding finite elements associated with the space U_h. Let the function $f \in C(\overline{\Omega}) \cap H^1(\Omega)$ and its partial derivatives $f_{x_i} \in L_\infty(\Omega)$. Let the function p satisfies conditions (13). Then

$$\left\|f - f^I\right\|_{C(\overline{\Omega})} \leq C \|\nabla f\|_{L_\infty(\Omega)} h, \tag{15}$$

$$\left\|p - p^I\right\|_{H^1(Q_T)} \leq C \|p\|_{H^2(Q_T)} (h + \tau). \tag{16}$$

Estimate (15) follows from the formula 76.3 in [15]. Estimate (16) follows from Theorem 3.2.1 in [13] and embedding theorem of $H^2(\Omega)$ in $C(\overline{\Omega})$ since $\Omega \in \mathbb{R}^3$ (the same is true for $\mathbb{R}^2$). The mesh regularity assumption is not necessary for (15), unlike (16).

Let the function $\varphi \in H_\varphi^2(Q_T)$. To solve the problem of the minimization of the functional (8), we introduce the Lagrangian

$$L(v) = E(u,c) - \int\limits_{Q_T} c(x) u_t \varphi_t dxdt + \int\limits_{Q_T} \nabla u \nabla \varphi dxdt - \int\limits_{S_T} p\varphi d\sigma dt, v = (u, \varphi, c). \tag{17}$$

By (9) and (14) the sum of integral terms in (17) equals zero. Thus, $L(v) = E(u,c)$. In other words, the addition of these terms to $E(u,c)$, does not change the Tikhonov functional. The reason of considering the Lagrangian instead of $E(u,c)$ is that it is easier (in certain sense) to find a stationary point of $L(v)$ compared with $E(u,c)$. To minimize the Lagrangian, we need to calculate its Frechet derivative and to set it to zero. Note that the function u depends on the coefficient c. In addition, below we will impose a constraint on the function φ requiring it to be the solution of the so-called "adjoint problem" (20). The latter means that φ also depends on c. Hence, in order to calculate the Frechet derivative rigorously, one should assume that variations of functions u and φ depend on variations of the coefficient c and calculate the Frechet derivative of $\widetilde{L}(c) := L(v(c))$. To do this, one needs, therefore, to consider Frechet derivatives of u, φ with respect to c in respectively defined functional spaces. However, this way, although completely rigorous, is anticipated to be quite space consuming, and we are not aware about previous publications where this way would be fully carried out for a CIP, although see [18] for an inverse problem of determining a boundary condition of a parabolic PDE; the latter is linear, unlike our CIP. We will consider the rigorous way in a forthcoming publication. At this point, however, following previous publications [2, 4, 9, 10, 11], we adopt a simpler heuristic the so-called "one shot" approach. Namely, we assume that in (12) functions u, φ, c can be varied independently on each other. Furthermore, whenever we discuss Frechet derivatives of L, we always mean mutually independent variations of all three components of the vector function v. However, as soon as this derivative is calculated, we assume that solutions u and φ of state (9) and adjoint (20) problems do depend on the coefficient c. The computational experience of both current and previous publications [4, 9, 10, 11] shows that this is sufficient.

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

$$L'(v)(\overline{v}) = 0, \quad \forall \overline{v} = (\overline{u}, \overline{\varphi}, \overline{c}) \in \bar{U} \tag{18}$$

where $L'(v)(\cdot)$ is the Frechet derivative of L at the point v under the above assumption of mutual independence of functions u, φ, c. To find $L'(v)(\overline{v})$, consider $L(v + \overline{v}) - L(v)$ $\forall \overline{v} \in \bar{U}$ and single out the linear, with respect to $\overline{v}$, part of this expression. Hence, we obtain from (17) and (18)

$$
\begin{aligned}
L'(v)(\overline{v}) = & \int_{\Omega} \left(\gamma (c - c_0) - \int_0^T u_t \varphi_t dt \right) \overline{c} dx \\
& + \left[\int_{Q_T} (-c u_t \overline{\varphi}_t + \nabla u \nabla \overline{\varphi}) \, dx dt - \int_{S_T} p \overline{\varphi} d\sigma dt \right] \\
& + \left[\int_{Q_T} (-c \varphi_t \overline{u}_t + \nabla \varphi \nabla \overline{u}) \, dx dt - \int_{S_T} \zeta_{\varepsilon_1} (g - u \mid_{S_T}) \overline{u} d\sigma dt \right] \\
& = 0, \forall \overline{v} = (\overline{u}, \overline{\varphi}, \overline{c}) \in \bar{U}.
\end{aligned} \tag{19}
$$

The term in the second line of (19) equals zero because of (9) and (14). To ensure that the term in the third line of (19) is zero, we assume first that there exists a function G satisfying (11). Next we set that the function φ is the solution of the following adjoint problem

$$
\begin{aligned}
c\varphi_{tt} - \triangle \varphi = 0, (x,t) \in Q_T, \\
\varphi(x,T) = \varphi_t(x,T) = 0, \\
\partial_n \varphi \mid_{S_T} = \zeta_{\varepsilon_1}(t) (g - u)(x,t), (x,t) \in S_T.
\end{aligned} \tag{20}
$$

Consider the function $\Phi(x,t) = \varphi(x,t) - \zeta_{\varepsilon_1}(t) [G(x,t) - a(x) u(x,t)]$. Then

$$
\begin{aligned}
c\Phi_{tt} - \Delta \Phi = \left[2\zeta_{\varepsilon_1} \nabla a \nabla u - 2ca \partial_t \zeta_{\varepsilon_1} u_t - \left(c\partial_t^2 - \Delta \right) (\zeta_{\varepsilon_1} G) \right], \\
\Phi(x,T) = \Phi_t(x,T) = 0, \quad \partial_n \Phi \mid_{S_T} = 0.
\end{aligned} \tag{21}
$$

Hence, there exists unique solution $\Phi \in H^2(Q_T)$ of the problem (21) and Φ satisfies condition (13). Therefore, there exists unique solution φ of the problem (20), and (13) holds for the function φ. The adjoint problem (20) should be solved backwards in time. For any function c satisfying (10) denote $u(c)$ and $\varphi(c)$ solutions of problems (9) and (20) respectively, both functions satisfy (13). Finally, to ensure that the first line of (19) equals zero, we set

$$
\gamma(c - c_0) - \int_0^T u_t \varphi_t \, dt = 0, x \in \Omega. \tag{22}
$$

Hence, it follows from (22) that in order to find the stationary point of the Lagrangian, we need to arrange an iterative procedure to approximate such a function $c(x)$, which would satisfy condition (10) and would be a solution of equation (22), where functions u and φ are solutions of state (9) and adjoint (20) problems respectively. The following lemma follows immediately from (14).

Lemma 3.1. *Consider an arbitrary function $c(x)$ satisfying condition (10) and assume that conditions (11) and (12) hold. Let functions $u, \varphi \in H^2(Q_T)$ be solutions of state (9) and adjoint (20) problems and $v = (u(c), \varphi(c), c)$ (i.e., v is not necesseraly a minimizer of the Lagrangian). Then*

$$L'(v)(\overline{v}) = \int_{\Omega} \left(\gamma (c - c_0) - \int_0^T u_t \varphi_t dt \right) \overline{c} dx, \forall \overline{v} = (\overline{u}, \overline{\varphi}, \overline{c}) \in \bar{U}.$$

3.2. A posteriori error estimate for the Lagrangian

Let the function $c^*(x)$ satisfying (3), (4) be the exact solution of our CIP, $g^*(x,t)$ be the corresponding function (5), and $u(c^*)$ be the solution of the Cauchy problem (1), (2) with $c := c^*$. Hence, $g^* - u^* \mid_{S_T} = 0$, meaning that the corresponding solution of the adjoint problem (20) $\varphi(c^*) = 0$. Denote $v^* = (u(c^*), 0, c^*) \in U$. Since the second stage of our two-stage procedure, the adaptivity, is a locally convergent numerical method, which takes a good approximation obtained on globally convergent first stage as a starting point, we work in this section in a small neighbourhood of the exact solution v^*. So, since $U \subset \bar{U}$ as a set, we work in section 3 in the set $V_\delta \subset \bar{U}$,

$$V_\delta = \left\{ \widehat{v} \in \bar{U} : \|\widehat{v} - v^*\|_{\bar{U}} < \delta \right\}, \tag{23}$$

where $\delta \in (0,1)$ is a sufficiently small number. In particular, δ can be linked with the parameter η of the global convergence theorem of [5], although we are not doing this here for brevity. Suppose that there exists a minimizer $v = (u(c), \varphi(c), c) \in U \cap \left(\overline{V}_\delta \diagdown \partial \overline{V}_\delta\right)$ of the Lagrangian L (17) satisfying (18) (and therefore (19)), and the function c satisfies condition (10). Note that because of an error in the data g in (5), it is not necessary that $v = v^*$. Assume that there exists a minimizer $v_h = (u_h(c_h), \varphi_h(c_h), c_h) \in U_h \cap \left(\overline{V}_\delta \diagdown \partial \overline{V}_\delta\right)$ of L on the discrete subspace U_h where the function c_h satisfies condition (10). Here and below $u_h(c_h) \in W_h^u$ and $\varphi_h(c_h) \in W_h^\varphi$ are finite element solutions of problems (9) and (20), respectively, with $c := c_h$ and under the assumption that boundary functions $p, g, u \mid_{S_T}$ in (9) and

(20) are the same as ones for functions $u(c), \varphi(c)$. The case when these boundary functions are approximated via finite elements can be considered along the same lines, and we are not doing this here for brevity. Hence, v_h is a solution of the following problem

$$L'(v_h)(\overline{v}) = 0, \forall \overline{v} \in U_h. \tag{24}$$

We now obtain a posteriori error estimate for the error in the Lagrangian. We have

$$L(v) - L(v_h) = \int_0^1 L'(\theta v + (1-\theta) v_h)\, d\theta = L'(v_h)(v - v_h) + R, \tag{25}$$

where the remainder term R is the second order of smallness with respect to δ. We ignore R, and the computational experience of both current and previous publications [9, 10, 11] shows that ignoring R does not have a visible impact on numerical results. Let $v^I = \left(u^I, \varphi^I, c^I\right)$ be the interpolant of the vector function v by finite elements of U_h. We have

$$v - v_h = (v^I - v_h) + (v - v^I). \tag{26}$$

Use the Galerkin orthogonality principle. Namely, by (24) and (26)

$$L'(v_h)(v - v_h) = L'(v_h)(v^I - v_h) + L'(v_h)(v - v^I) = L'(v_h)(v - v^I). \tag{27}$$

Hence (25) implies that the following approximate error estimate for the Lagrangian holds

$$L(v) - L(v_h) \approx L'(v_h)(v - v^I). \tag{28}$$

In (28) $v - v^I$ appear as interpolation errors. Hence, (15, 16) imply that one can estimate $v - v^I$ in terms of derivatives of v and the maximal grid step sizes h in space and τ in time, and this would specify the estimate (28).

If both state and adjoint problems are solved exactly, then Lemma 3.1 ensures that only the first line in the right hand side of (19) should be considered in a posteriori error analysis for the Lagrangian, and two other lines equal zero. In practice, however these two lines are not necessarily zeros because state and adjoint problems are solved by the FEM approximately. Hence, they should be taken into account in a posteriori error estimates. Consider first an "ideal" case when state and adjoint problems are solved exactly for the coefficient belonging to the discrete space V_h. Consider the space $\widetilde{U} = H_u^1(Q_T) \times H_\varphi^1(Q_T) \times V_h$. Consider a vector function $y_h :=$

$(u(c_h), \varphi(c_h), c_h) \in \widetilde{U} \cap (\overline{V}_\delta \diagdown \partial \overline{V}_\delta)$, where functions $u(c_h) := u(c_h, x, t) \in H^2(Q_T)$ and $\varphi(c_h) := \varphi(c_h, x, t) \in H^2(Q_T)$ are exact solutions of state (9) and adjoint (20) problems respectively with the function $c := c_h$ satisfying (10). By (14), (19) and Lemma 3.1

$$L'(y_h)(\overline{v}) = \int_\Omega \left[\gamma(c_h - c_0) - \int_0^T (u_t(c_h)\varphi_t(c_h))(x,t)\,dt \right] \overline{c} dx, \quad \forall \overline{v} = (\overline{u}, \overline{\varphi}, \overline{c}) \in \widetilde{U}. \tag{29}$$

Thus, we obtain

Theorem 3.1. *Assume that conditions (11) and (12) hold. Let the vector function* $v = (u(c), \varphi(c), c) \in U \cap (\overline{V}_\delta \diagdown \partial \overline{V}_\delta)$ *satisfies (18) and the vector function* $y_h = (u(c_h), \varphi(c_h), c_h) \in (\overline{V}_\delta \diagdown \partial \overline{V}_\delta)$ *be a minimizer of the Lagrangian* L *on the space* $\widetilde{U}$*. Let functions* c, c_h *satisfy condition (10). Then the following approximate a posteriori error estimate is valid*

$$\begin{aligned} |L(v) - L(y_h)| &\approx |L'(y_h)(v - v^I)| \\ &\leq C \|\nabla c\|_{L_\infty(\Omega)} h \left[\gamma \max_{\overline{\Omega}} |c_h - c_0| + \max_{\overline{\Omega}} \int_0^T (|u_t(c_h)| \cdot |\varphi_t(c_h)|)(x,t)\,dt \right]. \end{aligned} \tag{30}$$

Proof. Since y_h is a minimizer on the space $\widetilde{U}$, then $L'(y_h)(\overline{v}) = 0, \forall \overline{v} \in \widetilde{U}$. Since $v^I - v_h \in \widetilde{U}$, then $L'(y_h)(v^I - v_h) = 0$. Hence, it follows from (14), the definition of functions $u(c_h), \varphi(c_h)$ and (26), (27) that the following analog of (28) is valid $L(v) - L(y_h) \approx L'(y_h)(v - v^I)$. By (10) and (15) $\|c - c^I\|_{C(\overline{\Omega})} \leq Ch \|\nabla c\|_{L_\infty(\overline{\Omega})}$. The rest of the proof follows from (29)), where $\overline{c}$ should be replaced with $c - c^I$. $\square$

Remark 3.1. The estimate (30) indicates that the error in the Lagrangian can be decreased by refining the grid locally in those regions, where values of the function $B_h(x)$,

$$B_h(x) = \gamma |c_h - c_0|(x) + \int_0^T (|\partial_t u_h| \cdot |\partial_t \varphi_h|)(x,t)\,dt \tag{31}$$

are close to its maximal value. The latter forms the basis for the adaptivity technique, see subsection 6.4.

While it was assumed in Theorem 3.1 that state and adjoint problems with $c := c_h$ are solved precisely, in the next theorem we assume that they are solved via the FEM with a small error ε, see, e.g. [10] for some specific error estimates for the FEM for a second order hyperbolic PDE. It is natural to assume that

$$\left\| u\left(c_h\right) - u_h\left(c_h\right)\right\|_{H^1(Q_T)} \leq \varepsilon,\ \ \left\| \varphi\left(c_h\right) - \varphi_h\left(c_h\right)\right\|_{H^1(Q_T)} \leq \varepsilon. \tag{32}$$

Theorem 3.2. *Let conditions (11), (12) hold. Let vector functions* $v = \left(u\left(c\right), \varphi\left(c\right), c\right) \in U \cap \left(\overline{V}_\delta \diagdown \partial \overline{V}_\delta\right)$ *and* $v_h = \left(u_h\left(c_h\right), \varphi_h\left(c_h\right), c_h\right) \in U_h \cap \left(\overline{V}_\delta \diagdown \partial \overline{V}_\delta\right)$ *satisfy (18) and (24) respectively and coefficients* c, c_h *satisfy (10). Let* $\varepsilon \in (0,1)$ *be a sufficiently small positive number. Suppose that one can choose maximal grid step sizes in space and time* $h = h\left(\varepsilon\right)$ *and* $\tau = \tau\left(\varepsilon\right)$ *so small that the estimate (32) holds. Assume also that finite elements in* Ω *are regular. Then the following approximate a posteriori error estimate for the Lagrangian is valid*

$$\begin{aligned} &|L(v) - L(v_h)| \approx |L'(v_h)(v - v_h^I)| \\ &\quad \leq C \left\| \nabla c \right\|_{L_\infty(\overline{\Omega})} h \left[\gamma \max_{\overline{\Omega}} \left| c_h - c_0 \right| + \max_{\overline{\Omega}} \int_0^T \left(\left| u_t\left(c_h\right)\right| \cdot \left|\varphi_t\left(c_h\right)\right|\right)(x,t)\, dt \right] \\ &\quad + C\left(1 + 2d_2\right) \varepsilon \left(h + \tau\right) \left[\left\| u\left(c\right)\right\|_{H^2(Q_T)} + \left\| \varphi\left(c\right)\right\|_{H^2(Q_T)} \right]. \end{aligned} \tag{33}$$

Further, suppose that a priori estimate for the gradient of the unknown coefficient is $\left\| \nabla c \right\|_{L_\infty(\overline{\Omega})} \leq Z$*, where the positive constant* Z *is given (by the Tikhonov concept for ill-posed problems). Then with a constant* $C_1 = C_1\left(d_1, d_2, Z, \left\|P\right\|_{H^3(Q_T)}, \left\|G\right\|_{H^3(Q_T)}, Q_T\right) > 0$

$$\begin{aligned} &|L(v) - L(v_h)| \approx \left|L'(v_h)(v - v^I)\right| \\ &\quad \leq CZh \left[\gamma \max_{\overline{\Omega}} \left| c_h - c_0 \right| + \max_{\overline{\Omega}} \int_0^T \left(\left| u_t\left(c_h\right)\right| \cdot \left|\varphi_t\left(c_h\right)\right|\right)(x,t)\, dt \right] \\ &\quad + C_1 \varepsilon \left(h + \tau\right). \end{aligned} \tag{34}$$

Proof. Denote for brevity $u_h = u_h\left(c_h\right), \psi = \varphi\left(c\right) - \varphi^I\left(c\right).$ Since the third line of (19) can be estimated similarly with the second line, we consider only the second line with $u := u_h$. By (28) we should replace $\overline{\varphi}$ with ψ there. Denote

$$A_h = \int\limits_{Q_T} \left(-c_h u_{ht} \psi_t + \nabla u_h \nabla \psi\right) dxdt - \int\limits_{S_T} p\psi d\sigma dt. \tag{35}$$

Now, since the function $\psi \in H^1_\varphi(Q_T)$, then $\psi(x,T) = 0$. By (14)

$$\int\limits_{Q_T} \left(-c_h u_t(c_h)\psi_t + \nabla u(c_h)\nabla\psi\right) dxdt - \int\limits_{S_T} p\psi d\sigma dt = 0. \tag{36}$$

Since the function $\varphi(c)$ satisfies condition (13), then this implies in turn the estimate (16) for $\psi = \varphi(c) - \varphi^I(c)$. Hence, $\|\psi\|_{H^1(Q_T)} \leq C(h+\tau)\|\varphi(c)\|_{H^2(Q_T)}$. Thus, subtracting (36) from (35), we obtain

$$\begin{aligned} A_h &= \int\limits_{Q_T} \left(-c_h\left(u_h - u(c_h)\right)_t \psi_t + \nabla\left(u_h - u(c_h)\right)\nabla\psi\right) dxdt, \\ |A_h| &\leq (1+2d_2)\|u_h - u(c_h)\|_{H^1(Q_T)} \|\psi\|_{H^1(Q_T)} \\ &\leq (1+2d_2)\varepsilon\|\psi\|_{H^1(Q_T)} \leq C(1+2d_2)\varepsilon(h+\tau)\|\varphi(c)\|_{H^2(Q_T)}. \end{aligned} \tag{37}$$

This estimate for $|A_h|$ proves (33). To prove (34), we need to obtain upper estimates for norms $\|u(c)\|_{H^2(Q_T)}, \|\varphi(c)\|_{H^2(Q_T)}$. Consider the function $w = u(c) - P$. Then (9) and (11) imply that this function is the solution of the following initial boundary value problem

$$cw_{tt} = \Delta w - \left(c\partial_t^2 - \Delta\right)P,$$
$$w(x,0) = w_t(x,0) = 0, \partial_n w \mid_{S_T} = 0.$$

By (11) the function $(c\partial_t^2 - \Delta)P \in H^1(Q_T)$. Hence, we obtain similarly with (13) $w \in H^2(Q_T), \|w\|_{H^2(Q_T)} \leq C_1'\|P\|_{H^3(Q_T)}$, with a constant $C_1' = C_1'(d_1, d_2, Z, Q_T) > 0$. Hence, $\|u(c)\|_{H^2(Q_T)} \leq C_1$. The proof of the estimate $\|\varphi(c)\|_{H^2(Q_T)} \leq C_1$ can be obtained similarly via considering the function Φ in (21). □

Remark 3.2. Under a natural assumption $\varepsilon(h+\tau) << h$ (33),(34) indicate that one can approximately drop the third line in each of these estimates. In other words, Theorem 3.2 basically says that one can ignore terms in second and third lines of (19) when conducting a posteriori error analysis of the Lagrangian, provided that both state and adjoint problems are solved by the FEM with a good accuracy with $c := c_h$, although *not exactly*. The same is true for Theorems 3.3 and 3.4 in subsection 3.3.

3.3. A posteriori error estimate for the regularized unknown coefficient

Suppose that there exist two vector functions $v = (u(c), \varphi(c), c) \in U \cap \left(\overline{V}_\delta \diagdown \partial\overline{V}_\delta\right)$ and $v_h = (u_h(c_h), \varphi_h(c_h), c_h) \in U_h \cap \left(\overline{V}_\delta \diagdown \partial\overline{V}_\delta\right)$ satisfying conditions of Theorem

3.2. Denote $((\cdot,\cdot))$ the inner product in $\bar{U}^1$ and $[\cdot]$ the norm generated by this product. Let $L''(v_h)(\bar{v},w), \bar{v},w\in\bar{U}$ be the second Frechet derivative of the Lagrangian L at the point v_h, i.e. the Hessian. Consider a function $\psi\in\bar{U}$ and consider a solution $\widetilde{v}_\psi$ of the following problem, which we call the "Hessian problem",

$$\begin{aligned} -L''(v_h)(\overline{v},\widetilde{v}_\psi) &= ((\psi,\overline{v}))\quad \forall\overline{v}\in U_h,\\ \widetilde{v}_\psi &\in U\cap\left(\overline{V}_\delta\diagdown\partial\overline{V}_\delta\right). \end{aligned} \tag{38}$$

Assume that a solution $\widetilde{v}_\psi=(\widetilde{u}_\psi,\widetilde{\varphi}_\psi,\widetilde{c}_\psi)$ of this problem exists, and the function $\widetilde{c}_\psi$ satisfies (10). In (38) choose $\overline{v}=v-v_h$. Since by (18) $L'(v)(\widetilde{v}_\psi)=0$, we obtain

$$\begin{aligned} ((\psi,v-v_h)) &= -L''(v_h)(v-v_h,\widetilde{v}_\psi)\\ &= -L'(v)(\widetilde{v}_\psi)+L'(v_h)(\widetilde{v}_\psi)+R = L'(v_h)(\widetilde{v}_\psi)+R, \end{aligned} \tag{39}$$

where again R is the remainder term of the second order of smallness with respect to the parameter δ in (23). Thus, ignoring R, we obtain

$$((\psi,v-v_h)) = L'(v_h)(\widetilde{v}_\psi). \tag{40}$$

Because of dropping the term R, actually one should have "$\approx$" instead of "$=$" in (40), and this is why error estimates below are approximate ones. The formula (40) is the *main* factor enabling us to obtain approximate a posteriori error estimate for the regularized unknown coefficient.

The formulas (38), (39) and (40) where obtained in [11], although only a single function ψ was used there, i.e. the follow up analysis with functions ψ_k was not a part of [11].

Theorem 3.3. *Let conditions (11) and (12) hold. Suppose that there exist two vector functions* $v=(u(c),\varphi(c),c)\in U\cap\left(\overline{V}_\delta\diagdown\partial\overline{V}_\delta\right)$ *and* $v_h=(u_h(c_h),\varphi_h(c_h),c_h)\in U_h\cap\left(\overline{V}_\delta\diagdown\partial\overline{V}_\delta\right)$ *satisfying corresponding conditions of Theorem 3.2. Let* $P_h:\bar{U}^1\to U_h$ *and* $Q_h:L_2(\Omega)\to V_h$ *be orthogonal projection operators of spaces* $\bar{U}^1$ *and* $L_2(\Omega)$ *on their respective subspaces* U_h *and* V_h*. Let* $\{\psi_k\}_{k=1}^M\subset U_h$ *be an orthonormal basis in the space* U_h*. Suppose that for each vector function* ψ_k *there exists a solution* $\widetilde{v}_{\psi_k}=(\widetilde{u}_{\psi_k},\widetilde{\varphi}_{\psi_k},\widetilde{c}_{\psi_k})$ *of the problem (38) with the function* $\widetilde{c}_{\psi_k}$ *satisfying condition (10). Then the following approximate a posteriori* *error estimate for the target coefficient is valid*

$$\|Q_h c - c_h\|_{L_2(\Omega)} \leq [P_h v - v_h] = \left[\sum_{k=1}^{M} |((\psi_k, P_h v - v_h))|^2 \right]^{1/2} \leq \left[\sum_{k=1}^{M} \left| L'(v_h)(\widetilde{v}_{\psi_k} - \widetilde{v}_{\psi_k}^I) \right|^2 \right]^{1/2}. \tag{41}$$

In particular, assume that problems (9) and (20) are solved exactly for $c := c_h$. *Then*

$$\|Q_h c - c_h\|_{L_2(\Omega)} \leq$$
$$\sqrt{M} C \max_k \|\nabla \widetilde{c}_{\psi_k}\|_{C(\overline{\Omega})} h \left[\gamma \max_{\overline{\Omega}} |c_h - c_0| + \max_{\overline{\Omega}} \int_0^T \left(|u_t(c_h)| \cdot |\varphi_t(c_h)| \right)(x,t)\, dt \right], \tag{42}$$

where $M = \dim(U_h)$.

Consider now the case when problems (9) and (20) with $c := c_h$ *are solved approximately by the FEM, i.e., assume that (32) holds. Also, let finite elements in* Ω *be regular. Then*

$$\|Q_h c - c_h\|_{L_2(\Omega)} \leq$$
$$\sqrt{M} C \max_k \|\nabla \widetilde{c}_{\psi_k}\|_{C(\overline{\Omega})} h \left[\gamma \max_{\overline{\Omega}} |c_h - c_0| + \max_{\overline{\Omega}} \int_0^T \left(|u_t(c_h)| \cdot |\varphi_t(c_h)| \right)(x,t)\, dt \right]$$
$$+ \sqrt{M} C (1 + 2d_2)\, \varepsilon\, (h + \tau) \left[\max_k \|u(\widetilde{c}_{\psi_k})\|_{H^2(Q_T)} + \max_k \|\varphi(\widetilde{c}_{\psi_k})\|_{H^2(Q_T)} \right]. \tag{43}$$

Proof. We have $w = P_h w + (I - P_h) w, \forall w \in \bar{U}^1$. Since the vector $(I - P_h) w$ is orthogonal to the subspace U_h, then $((\psi_k, (I - P_h) w)) = 0, \forall w \in \bar{U}^1, k = 1, ..., M$. Hence, $((\psi_k, v)) = ((\psi_k, P_h v)) + ((\psi_k, (I - P_h) v)) = ((\psi_k, P_h v))$. Hence, $((\psi_k, v - v_h)) = ((\psi_k, P_h v - v_h))$. Therefore, (40) implies that

$$((\psi_k, P_h v - v_h)) = L'(v_h)(\widetilde{v}_{\psi_k}). \tag{44}$$

Let $\widetilde{v}_{\psi_k}^I$ be the interpolant of the vector function $\widetilde{v}_{\psi_k}$ by finite elements of the space U_h. Then by (24) $L'(v_h)\left(\widetilde{v}_{\psi_k}^I\right) = 0$. Hence, using an analog of (26), we obtain $L'(v_h)(\widetilde{v}_{\psi_k}) = L'(v_h)\left(\widetilde{v}_{\psi_k} - \widetilde{v}_{\psi_k}^I\right) + L'(v_h)\left(\widetilde{v}_{\psi_k}^I\right) = L'(v_h)\left(\widetilde{v}_{\psi_k} - \widetilde{v}_{\psi_k}^I\right)$. Hence, by (44)

$$((\psi_k, P_h v - v_h)) = L'(v_h)\left(\widetilde{v}_{\psi_k} - \widetilde{v}_{\psi_k}^I\right). \tag{45}$$

Since $\|Q_hc - c_h\|^2_{L_2(\Omega)} \leq [P_hv - v_h]^2$, then by (45)

$$\|Q_hc - c_h\|^2_{L_2(\Omega)} \leq [P_hv - v_h]^2 = \sum_{k=1}^{M} |((\psi_k, P_hv - v_h))|^2 = \sum_{k=1}^{M} \left| L'(v_h)\left(\widetilde{v}_{\psi_k} - \widetilde{v}^I_{\psi_k}\right)\right|^2 \tag{46}$$

Thus, (41) follows from (46). Estimate (42) follows immediately from (41) and Theorem 3.1. Estimate (43) follows from (42), (46) and Theorem 3.2. □

Remark 3.3. Note that the right hand sides of a posteriori error estimates (42) and (43) for the regularized unknown coefficient have basically the same form as ones for the accuracy of the Lagrangian in Theorems 3.1 and 3.2, respectively. This is convenient for computations. Thus, refining mesh, as in Remark 3.1, one might improve the accuracy of the reconstruction of both the Lagrangian and the regularized coefficient. Numerical studies of [11] seem to indicate that required solutions of the Hessian problem exist. An inconvenient point of estimates (42), (43) is that one should estimate maximal values depending on functions $\widetilde{v}_{\psi_k}$. To mitigate this, we impose a little bit more stringent condition in Theorem 3.4.

Theorem 3.4. *Assume that (38) in one of conditions of Theorem 3.3 is replaced with*

$$\widetilde{v}_{\psi_k} \in \{\widehat{v} \in U : \|\widehat{v} - v^*\|_U < \delta < 1\}, \tag{47}$$

which means that in (23) the space $\bar{U}$ is replaced with the space U with a stronger norm. Let the rest of conditions of Theorem 3.3 holds. Assume that the exact unknown coefficient $c^*(x)$ satisfies (3), (4) and a priori estimate for its gradient is $\|\nabla c^*\|_{C(\overline{\Omega})} \leq Z$, where the positive constant Z is known (by the Tikhonov concept for ill-posed problems). Let the function $u^*(x,t)$ be the solution of the Cauchy problem (1), (2) with $c := c^*$in the case when in (2) $\delta(x - x_0)$ is replaced with a non-zero function $F(x) \in C^\infty(\mathbb{R}^3)$ with a compact support and such that $F(x) = 0$ in $\overline{\Omega}$. Then with a constant $C_2 = C_2(d_1, d_2, Z, Q_T, F) > 0$ the following estimate is valid

$$\begin{aligned} \|Q_hc - c_h\|_{L_2(\Omega)} &\leq \\ &\leq \sqrt{M}CZh\left[\gamma \max_{\overline{\Omega}} |c_h - c_0| + \max_{\overline{\Omega}} \int_0^T (|u_t(c_h)| \cdot |\varphi_t(c_h)|)(x,t)\,dt\right] \\ &+ \sqrt{M}C_2\varepsilon(h + \tau). \end{aligned} \tag{48}$$

Proof. By (47) $\|u(\widetilde{c}_{\psi_k})\|_{H^2(Q_T)} \leq \|u^*\|_{H^2(Q_T)} + \delta$. Since it was observed above that $\varphi(c^*) = 0$, then (47) leads to $\|\varphi(\widetilde{c}_{\psi_k})\|_{H^2(Q_T)} \leq \delta$. Hence, it follows from

(43) that in order to prove (48), it is sufficient to estimate from the above the norm $\|u^*\|_{H^2(Q_T)}$. Since the function $F(x)$ has a compact support, then, as it was established in §2 of Chapter 4 of [22], it follows from the finite speed of propagation property for hyperbolic equations that $\|u^*\|_{H^2(Q_T)} \leq C_2$. $\square$

Remark 3.4. Although the number M is large for small h, still Theorems 3.3 and 3.4 show that the error in the regularized coefficient is basically determined by the value of the gradient with respect to this coefficient. In other words, the mesh refinement recommended in Remark 3.1 should likely decrease the error not only in the Lagrangian but in the regularized target coefficient as well, and we observe this in our computations [8]. In the future we hope to improve these error estimates in such a way that the number M would not be present in them.

3.4. A general framework for derivation of analogs of theorems 3.1-3.4 for different types of CIPs

We now outline a general framework of derivations of a posteriori error estimates like ones in Theorems 3.1-3.4 for CIPs for three main types of PDEs of the second order: hyperbolic, parabolic and elliptic. Suppose we have a CIP for one of these three types of PDEs and that we want to apply the adaptivity technique, which is similar to the one described above. Then we propose the following framework:

Step 1. Write down the Tikhonov functional similar with (8) for hyperbolic equations, then write the Lagrangian similarly to (17).

Step 2. Derive the Frechet derivative of the Lagrangian, assuming that solutions of state and adjoint problems are independent on the unknown coefficient.

Step 3. Using the definition of the weak H^1 solution of the original PDE, make sure that integral terms, which are not responsible for the unknown coefficient, equal zero similarly with Lemma 3.1.

Step 4. Similarly with Theorem 3.1 derive a posteriori error estimate for the Lagrangian, assuming that state and adjoint problems are solved exactly.

Step 5. Assuming that state and adjoint problems are solved approximately by the FEM, derive an analog of Theorem 3.2. To do so, introduce an obvious analog of the assumption (32). Next, subtract from corresponding integral terms of the Lagrangian integral identities which define weak solutions of state and adjoint problems, similarly with (35), (36), (37). Then using (32), Galerkin orthogonality and analogs of interpolation estimates (15), (16), one obtains an analog of Theorem 3.2, as well as an analog of Remark 3.1. These provide a recommendation for mesh refinement. In particular, they indicate that the impact of certain integral terms in the Frechet derivative of the Lagrangian is not essential compared with the Frechet derivative with respect to the unknown coefficient.

Step 6. To obtain a posteriori estimate for the regularized coefficient, observe that formulas (38)-(40) are general ones, which are valid for a general Lagrangian. Therefore, derivations of analogs of Theorems 3.3, 3.4 from an analog of Theorem 3.2 can be obtained straightforwardly from the above.

Acknowledgements

This work was supported by the U.S. Army Research Laboratory and U.S. Army Research Office grants W911NF-05-1-0378 and W911NF-08-1-0470. The authors are grateful to Prof. M.Yu. Kokurin for fruitful discussions.

References

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

[2] O.M. Alifanov, E. A. Artiukhin and S. V. Rumiantsev, *Extreme Methods for Solving Ill-Posed Problems With Applications to Inverse Heat Transfer Problems*, Begell House, New York, 1995.
order hyperbolic equations, SIAM J. Num. Anal., 13, 564-576

[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, available on-line at Chalmers Preprint Series ISSN 1652-9715, 2009:6 and at http://www.ma.utexas.edu/mp_arc/, 2009.

[7] L. Beilina and M. V. Klibanov, A globally convergent numerical method and the adaptivity technique for a hyperbolic coefficient inverse problem. Part I: analytical study, available on-line at http://www.ma.utexas.edu/mp_arc/, 2009.

[8] L. Beilina and M. V. Klibanov, A globally convergent numerical method and the adaptivity technique for a hyperbolic coefficient inverse problem. Part II: numerical studies, available on-line at http://www.ma.utexas.edu/mp_arc/, 2009.

[9] L. Beilina and C. Clason. An adaptive hybrid fem/fdm method for an inverse scattering problem in scanning acoustic microscopy. *SIAM J. Sci. Comp.*, 28(1):382 402, 2006.

[10] 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.

[11] 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.

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

[13] P.G. Ciarlet, *The Finite Element Method for Elliptic Problems*, SIAM, 2002.

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

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

[16] K. Eriksson, D. Estep and C. Johnson, *Computational Differential Equations*, Cambridge University Press and Studentlitteratur, ISBN 91-44-49311-8, 1996.

[17] L. C. Evans, *Partial Differential Equations, AMS, Providence,* AMS, Providence, RI, 1998.

[18] A. Hasanov, Simultaneous determination of source terms in a linear parabolic problem from final overdetermination: Weak solution approach, *J. Math. Anal. Appl.* 330, 766-799, 2007.

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

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

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

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

[23] M.M. Lavrentieve, V.G. Romanov and S.P. Shishatskii, *Ill-Posed Problems of Mathematical Physics and Analysis*, AMS, Providence, RI, 1986.

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

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