Carleman contraction mapping for a 1D inverse scattering problem with experimental time-dependent data

Thuy T Le¹, Michael V Klibanov¹,∗, Loc H Nguyen¹, Anders Sullivan² and Lam Nguyen²

¹ Department of Mathematics and Statistics, University of North Carolina at Charlotte, Charlotte, NC, 28223, United States of America
² US Army Research Laboratory, 2800 Powder Mill Road, Adelphi, MD 20783-1197, United States of America

E-mail: tle55@uncc.edu, mklibanv@uncc.edu, loc.nguyen@uncc.edu, anders.j.sullivan.civ@mail.mil and lam.h.nguyen2civ@mail.mil

Received 22 September 2021, revised 7 January 2022
Accepted for publication 1 February 2022
Published 24 February 2022

Abstract
It is demonstrated that the contraction mapping principle with the involvement of a Carleman weight function works for a coefficient inverse problem for a 1D hyperbolic equation. Using a Carleman estimate, the global convergence of the corresponding numerical method is established. Numerical studies for both computationally simulated and experimentally collected data are presented. The experimental part is concerned with the problem of computing dielectric constants of explosive-like targets in the standoff mode using severely underdetermined data.

Keywords: one-dimensional wave equation, Carleman estimate, iterative method, contraction principle, numerical solution, experimental data, global convergence

(Some figures may appear in colour only in the online journal)

1. Introduction

The phenomenon of multiple local minima and ravines of conventional least squares cost functionals for coefficient inverse problems (CIPs) is well known, see, e.g. [38] for a numerical example. Some cost functionals for various CIPs can be found in, e.g. [10, 13, 14]. Due to this phenomenon, convergence of a numerical method of the minimization of that functional to the true solution of the corresponding CIP can be guaranteed only if its starting point is

∗Author to whom any correspondence should be addressed.
located in a sufficiently small neighborhood of this solution. To avoid the latter, the so-called convexification method was initially proposed theoretically in [19, 20] and more recently this method was tested on a variety of CIPs. The corresponding results for multidimensional CIPs are summarized in the recently published book [25]. As to the various versions of the convexification method for the 1D CIP of this publication, we refer to publications [21, 22, 24, 39], where the same experimental data as ones discussed in this paper were treated. In [40] the same CIP, although, without experimental data was treated by another version of the convexification method.

In the convexification, a weighted Tikhonov-like functional $J_{\lambda, \beta}$ is constructed first, where $\lambda, \beta > 0$ are two parameters. The weight is the Carleman weight function (CWF). This is the function, which is used as the weight in the Carleman estimate for the corresponding partial differential operator, see, e.g. [5, 6, 17, 25, 30, 42] for some publications on Carleman estimates. That functional $J_{\lambda, \beta}$ is defined on a bounded convex set $S \subset H$, where $H$ is an appropriate Hilbert space. Next, it is proven that, for an appropriate choice of the parameters $\lambda, \beta$, the functional $J_{\lambda, \beta}$ is strictly convex on $S$, has unique minimizer on $S$ and minimizers generate a sequence, which converges to the true solution of the original CIP as long as the level of noise in the data tends to zero. Let $d(S)$ be the diameter of the set $S$. An important point here is that a smallness condition is not imposed on $d(S)$. This means that convexification is a globally convergent numerical method in terms of definition given below. We call the method outlined in this paragraph the ‘first generation of the convexification method’.

**Definition.** We call a numerical method for a CIP globally convergent if there is a theorem claiming that this method delivers at least one point in a sufficiently small neighborhood of the true solution of that CIP without an advanced knowledge of a small neighborhood of this solution.

The above functional $J_{\lambda, \beta}$ is not a quadratic one. The main new element of this paper is that we minimize a sequence of quadratic functionals. More precisely, unlike the above, we construct a sequence of linear boundary value problems (BVPs) for certain PDEs with overdetermined boundary conditions and non-local terms. To solve each of these BVPs, we apply a weighted version of the quasi-reversibility method (QRM). The weight is again the CWF. This is the reason why we call this method ‘Carleman quasi-reversibility method’ (CQRM). For each BVP, CQRM minimizes a quadratic weighted functional on a bounded set $S'$, which is an analog of the abovementioned set $S$. We analytically establish the key convergence estimate for the sequence of minimizers of these functionals. Our convergence estimate implies the global convergence of that sequence to the true solution of our CIP. We call the technique of this paper ‘the second generation of the convexification method’.

**Remark 1.1.** The convergence estimate mentioned in the previous paragraph is similar with the estimate of the classical contraction mapping principle, see the first item of remark 8.1 in section 8. This explains the title of our paper.

Furthermore, due to its connection with the contraction mapping, that convergence estimate implies a rapid convergence of our technique. As a result, computations for our experimental data are performed in real time here, also, see remark 11.1 in section 11 as well as section 12. On the other hand, real time computations are not claimed in the previous works of this research group on the first generation of the convexification method applied to the same CIP and with the same experimental data as ones used in this paper [21, 22, 24, 39]. The real time computations are obviously important for our described real world application described below in this section. Thus, the real time computations for experimental data present an important advantage of the technique of this paper over the first generation of the convexification method.
The QRM was first proposed in [29], also, see, e.g. [7, 8, 18, 25] and references cited therein for some follow up publications. Even though Carleman estimates were used in [7, 8, 18, 25] to establish convergence rates, the CWFs were not involved in numerical schemes.

For the first time, the second generation of the convexification method involving CQRM was published in [2]. We also refer to works [3, 4] of the same research group for their later publications. In these papers, globally convergent numerical methods for CIPs for hyperbolic PDEs in $\mathbb{R}^n$ were developed. It is assumed in [2–4] that one of initial conditions is not vanishing everywhere in the closed domain of interest. In other words, papers [2–4] work in the framework of the Bukhgeim–Klibanov method, see [9] for the originating work on this method as well as, e.g. [5, 6, 16, 17, 25, 42] for some follow up publications. The major difference between works [2–4] and all above cited publications of our research group on the convexification, including the current one, is that in our works either one of initial conditions is the $\delta$-function and another one is zero, or a similar requirement holds for the Helmholtz equation. The only exception is the paper [23], which works within the framework of the method of [9], also, see chapter 9 of [25] for the same result as the one in [23]. In [31] the second generation of the convexification was applied to solve an inverse problem of the determination of the initial condition of a quasilinear parabolic equation. We refer to papers [12, 27] for some globally convergent numerical methods for CIPs with the Dirichlet-to-Neumann map data. In these works the number $m$ of free variables in the data exceeds the number $n$ of free variables in the unknown coefficient, $m > n$. In our paper $m = n = 1$. Also, $m = n$ in all other above cited works on the convexification.

We consider in this paper a CIP for a 1D hyperbolic PDE. We show that our CIP has a direct application in the problem of the standoff detection and identification of antipersonnel land mines and improvised explosive devices (IEDs). Thus, in the computational part of this paper, we present results of the numerical performance of our technique for both computationally simulated and experimentally collected data for targets mimicking antipersonnel land mines and IEDs. The experimental data of this paper were collected by the forward looking radar of the US Army Research Laboratory [34]. Since these data were described in some previous publications of our research group [15, 21, 22, 24, 28, 39], then we do not describe them here.

From the applied standpoint, our goal is to compute approximate values of dielectric constants of above mentioned targets. We point out that our experimental data are severely under determined. Indeed, any target is a 3D object. On the other hand, we have only one experimentally measured time resolved curve for each target. Therefore, we can compute only a sort of an average value of the dielectric constant of each target. This is the reason of our mathematical modeling of our experimental data by a 1D hyperbolic PDE rather than by its 3D analog. We believe that our results for experimental data might potentially help to decrease the false alarm rate in the problem of the standoff detection and identification of antipersonnel land mines and IEDs.

There is a classical Gelfand–Levitan (GL) method [32] for solutions of 1D CIPs for some hyperbolic PDEs. This method does not rely on the optimization, and, therefore, avoids the phenomenon of local minima. It reduces the original CIP to a linear integral equation of the second kind. This is so-called ‘GL equation’. However, the questions of uniqueness and stability of the solution of GL for the case of noisy data are open, see, e.g. lemma 2.4 in the book [36, chapter 2]. This lemma is valid only in the case of noiseless data. However, the realistic data are always noisy. In addition, it was demonstrated numerically in [15] that GL cannot work well for exactly the same experimental data as the ones we use in the current paper. On
the other hand, it was demonstrated in [21, 22, 24, 39] that the first generation of the convexification method works well with these data. The same is true for the second generation of the convexification method of this paper.

Uniqueness and Lipschitz stability theorems of the CIP considered here are well known. Indeed, it was shown in, e.g., [39, 40] that, using a change of variables, one can reduce our CIP to a similar CIP for the equation \( v_{tt} = v_{yy} + r(y) v, \) \( y \in \mathbb{R} \) with the unknown coefficient \( r(y) \). We refer to [36, theorem 2.6 of chapter 2] for the Lipschitz stability estimate for the latter CIP. In addition, both uniqueness and Lipschitz stability results for our CIP actually follow from theorem 8.1 below as well as from the convergence analysis of [24] for the first generation of the convexification method for this CIP.

This paper is arranged as follows. In section 2 we state both forward and inverse problems. In section 3 we derive a nonlinear BVP with non local terms. In section 4 we describe our iterative solution of this BVP. In section 5 we formulate the Carleman estimate for the principal part of the PDE operator of that BVP. In section 6 we prove the strong convexity of a functional of section 5 on an appropriate bounded set in a Hilbert space. In section 7 we formulate two methods for finding the unique minimizer of that functional: the gradient descent method and the gradient projection method and prove the global convergence to the minimizer for each of them. In section 8 we establish the contraction mapping property and prove the global convergence of the method of section 4. In section 9 we formulate two more global convergence theorems, which follow from results of sections 7 and 8. Numerical results with simulated and experimental data are presented in sections 10 and 11 respectively. Concluding remarks are given in section 12.

2. Statements of forward and inverse problems

Below all functions are real valued ones. Let \( b > 1 \) be a known number, \( x \in \mathbb{R} \) be the spatial variable and the function \( c(x) \in C^3(\mathbb{R}) \), represents the spatially distributed dielectric constant. We assume that

\[
c(x) \in [1, b], \quad x \in \mathbb{R},
\]

\[
c(x) = 1 \quad \text{if} \quad x \in (-\infty, \varepsilon] \cup [1, \infty),
\]

where \( \varepsilon \in (0, 1) \) is a small number. Let \( T \) be a positive number. We consider the following Cauchy problem for a 1D hyperbolic PDE with a variable coefficient in the principal part of the operator:

\[
c(x) u_{tt} = u_{xx}, \quad x \in \mathbb{R}, \quad t \in (0, T),
\]

\[
u(x, 0) = 0, \quad u_t (x, 0) = \delta (x).
\]

The problem of finding the function \( u(x,t) \) from conditions (2.3) and (2.4) is our forward problem.

Let \( \tau (x) \) be the travel time needed for the wave to travel from the point source at \{0\} to the point \( x \),

\[
\tau (x) = \int_0^x \sqrt{c(s)}ds.
\]

By (2.5) the following 1D analog of the eikonal equation is valid:

\[
\tau' (x) = \sqrt{c(x)}.
\]
Let $H(z), z \in \mathbb{R}$ be the Heaviside function,

$$H(z) = \begin{cases} 1, & z > 0, \\ 0, & z < 0. \end{cases}$$

**Lemma 2.1** [24, lemma 2.1]. For $x \geq 0$, the function $u(x, t)$ has the form:

$$u(x, t) = H(t - \tau(x)) \left[ \frac{1}{2c^{1/4}(x)} + \hat{u}(x, t) \right],$$

where the function $\hat{u} \in C^2(t \geq \tau(x))$ and $\hat{u}(x, \tau(x)) = 0$. In particular,

$$\lim_{t \to \tau(x)} u(x, t) = \frac{1}{2c^{1/4}(x)}.$$

We also refer to books of Romanov [36, formulas (2.50) and (2.51)], [37, lemma 1.2.1] for results, which are similar with the one of lemma 2.1. Far more challenging similar results in the 3D case can also be found in these books, see [36, theorem 4.1], [37, lemma 2.2.1].

**Lemma 2.2** (Absorbing boundary conditions [24, 39]). Let $b > \varepsilon$ be the number in (2.1). Let $x_1 \geq b$ and $x_2 \leq \varepsilon$ be two arbitrary numbers. Then the solution $u(x, t)$ of forward problem (2.3) and (2.4) satisfies the absorbing boundary conditions at $x = x_1$ and $x = x_2$, i.e.

$$u_t(x_1, t) + u_t(x_1, t) = 0, \quad t \in (0, T),$$

$$u_t(x_2, t) - u_t(x_2, t) = 0, \quad t \in (0, T).$$

We are interested in the following inverse problem:

**Coefficient inverse problem (CIP).** Suppose that the following two functions $g_0(t)$, $g_1(t)$ are known:

$$u(\epsilon, t) = g_0(t), \quad u_4(\epsilon, t) = g_1(t), \quad t \in (0, T).$$

Determine the function $c(x)$ for $x \in (\varepsilon, 1)$, assuming that the number $b > \varepsilon$ in (2.1) is known.

**Remark 2.1.** Note that only the function $g_0(t)$ can be measured. As to the function $g_1(t)$, it follows from (2.10) that

$$g_1(t) = g_0'(t).$$

We differentiate noisy function using the Tikhonov regularization method [41]. Since this method is well known, we do not describe it here.

### 3. A boundary value problem for a nonlinear PDE with non-local terms

We now introduce a change of variable

$$q(x, t) = u(x, t + \tau(x)).$$

We will consider the function $q(x, t)$ only for $t \geq 0$. Using (2.6) and (3.1), we obtain

$$q_{xx} - 2q_{xx} \tau' - q_{xx}'' = 0.$$
Furthermore, it follows from (2.8)
\[
q(x, 0) = \frac{1}{2c^{1/4}(x)} \neq 0. \tag{3.3}
\]
By (2.5) and (3.3)
\[
\tau''(x) = -\frac{q_x(x, 0)}{2q^3(x, 0)} \tag{3.4}
\]
Substituting (2.6), (3.3) and (3.4) in (3.2), we obtain
\[
L(q) = q_{xx} - q_{xt} + \frac{1}{2} q^2 x_{0} + q_t q_x = 0. \tag{3.5}
\]
Equation (3.5) is a nonlinear PDE with respect to the function \(q(x, t)\) with nonlocal terms \(q_x(x, 0)\) and \(q(x, 0)\). We now need to obtain boundary conditions for the function \(q\).

By (2.2) and (2.5)
\[
\tau(x) = x \text{ for } x \in [0, \varepsilon].
\]

Hence, (2.11), (2.12) and (3.1) lead to
\[
q(\varepsilon, t) = g_0(t + \varepsilon), \quad q_x(\varepsilon, t) = 2g'_0(t + \varepsilon), \quad t \in (0, T). \tag{3.6}
\]

We will solve equation (3.5) in the rectangle
\[
\Omega = \{(x, t) : x \in (\varepsilon, b), t \in (0, T)\}. \tag{3.7}
\]

By (3.1)
\[
q_x(x, t) = u_x(x, t + \tau(x)) + u_t(x, t + \tau(x))\tau'(x). \tag{3.8}
\]

By (2.1), (2.2) and (2.5) \(\tau'(b) = 1\). Hence, using (2.9) and (3.8), we obtain
\[
q_x(b, t) = 0. \tag{3.9}
\]

It follows from (2.1) and (3.3) that
\[
\frac{1}{2b^{1/4}} \leq q(x, 0) \leq \frac{1}{2}, \quad x \in [\varepsilon, b]. \tag{3.10}
\]

In addition, we need \(q \in C^2(\overline{\Omega})\) and we also need to bound the norm \(\|q\|_{C^2(\overline{\Omega})}\) from the above.

Let \(R > 0\) be an arbitrary number. We define the set \(B(R, g_0)\) as
\[
B(R, g_0) = \begin{cases}
q \in H^4(\Omega) : \|q\|_{H^4(\Omega)} < R, \\
q(\varepsilon, t + \varepsilon) = g_0(t + \varepsilon), \quad q_x(\varepsilon, t) = 2g'_0(t + \varepsilon), \\
q_x(b, t) = 0, \\
\frac{1}{2b^{1/4}} \leq q(x, 0) \leq 1/2, \quad x \in [\varepsilon, b].
\end{cases} \tag{3.11}
\]

We assume below that
\[
B(R, g_0) \neq \emptyset. \tag{3.12}
\]

By embedding theorem \(B(R, g_0) \subset C^2(\overline{\Omega})\) and
\[
\|q\|_{C^2(\overline{\Omega})} \leq KR, \quad \forall q \in B(R, g_0). \tag{3.13}
\]
where the constant \( K = K (\Omega) > 0 \) depends only on the domain \( \Omega \). Using (3.10), we define the function \( q^0(x, 0) \) as:

\[
q^0(x, 0) = \begin{cases} 
q(x, 0) & \text{if } q(x, 0) \in \left[ 1/\left(2b^{1/4}\right), 1/2 \right], \\
1/\left(2b^{1/4}\right) & \text{if } q(x, 0) < 1/\left(2b^{1/4}\right), \\
1/2 & \text{if } q(x, 0) > 1/2.
\end{cases}
\]

\( \forall x \in [\varepsilon, b] \).

Then the function \( q^0(x, 0) \) is piecewise continuously differentiable in \([\varepsilon, b]\) and by (3.13) and (3.14)

\[
\left\| q^0(x, 0) \right\|_{C[\varepsilon, b]} \leq \max_{x \in [\varepsilon, b]} | q^0(x, 0) | \leq KR, \quad \forall q \in B(R, g_0), \quad \text{max}_{x \in [\varepsilon, b]} \left| q^0(x, 0) \right| \leq 1,
\]

(3.15)

\[
\frac{1}{2b^{1/4}} \leq q^0(x, 0) \leq \frac{1}{2}, \quad x \in [\varepsilon, b].
\]

(3.16)

Thus, (3.5), (3.6), (3.9), (3.14) and (3.16) result in the following BVP for a nonlinear PDE with non-local terms:

\[
q_{xx}(x, t) - q_{xt}(x, t) \frac{1}{2(q^0(x, 0))^2} + q_t(x, t) \frac{q_x(x, 0)}{2(q^0(x, 0))^3} = 0 \quad \text{in } \Omega,
\]

(3.17)

\[
q(\varepsilon, t) = g_0(t + \varepsilon), \quad q_x(\varepsilon, t) = 2g'_0(t + \varepsilon), q_x(b, t) = 0.
\]

(3.18)

Thus, we focus below on the numerical solution of the following problem.

**Problem 3.1.** Find a function \( q \in B(R, g_0) \) satisfying conditions (3.17) and (3.18), where the function \( q^0(x, 0) \) is defined in (3.14).

Suppose that we have solved problem 3.1. Then, using (3.3) and (3.14), we set

\[
c(x) = \frac{1}{\left(2q^0(x, 0)\right)^2}.
\]

(3.19)

4. Numerical method for problem 3.1

4.1. The function \( q_0(x, t) \)

We now find the first approximation \( q_0(x, t) \) for the function \( q(x, t) \). Using (2.2), we choose \( c(x) \equiv 1 \) as the first guess for the function \( c(x) \). Hence, by (3.3),

\[
q_0(x, 0) \equiv \frac{1}{2}.
\]

(4.1)

We now need to find the function \( q_0(x, t) \). To do this, drop the nonlinear third term in the left-hand side of equation (3.17) and, using (4.1) and (3.14), set \( 1/(2q^0(x, 0))^2 = 2 \). Then (3.17) and (3.18) become:
parameters will be chosen later. Assume that functionals $q_n$ becomes a linear on with respect to the function $\alpha$ where $e^{2\lambda \nu}$ is the CWF for the operator $\partial^2_x - 2\partial^2_{tt}$ [40]

$$e^{2\lambda \nu} = e^{-2\lambda (x + \alpha \beta)}$$

where $\alpha \in (0, 1/2)$ is the parameter, and $\beta \in (0, 1)$ is the regularization parameter. Both parameters will be chosen later.

Theorem 6.1 guarantees that, for appropriate values of parameters $\lambda, \beta$, there exists unique minimizer $q_{0, \text{min}} \in \overline{B(R, g_0)}$ of the functional $J^{(0)}_{\lambda, \beta}(q_0)$.

4.2. The function $q_n(x, t)$ for $n \geq 1$

Assume that functionals $J^{(0)}_{\lambda, \beta}(q_n) : \overline{B(R, g_0)} \to \mathbb{R}$ are defined for $m = 0, \ldots, n - 1$, and their minimizers functions $(q_{0, \text{min}}^{(m)} = q_{m, \text{min}})_{m=0}^{n-1} \subset \overline{B(R, g_0)}$ are constructed already, all for the same values of parameters $\lambda, \alpha, \beta$. Replace in (3.17) $q(x, t)$ with $q_n(x, t), q^0(x, 0)$ with $q^{(0-1, \text{min}}_{n-1}(x, 0), q_n(x, t)$ with $\partial_x q^{(0-1, \text{min}}_{n-1}(x, t)$ and $q_n(x, t)$ with $\partial q^{(0-1, \text{min}}_{n-1}(x, t)$. Then problem (3.17) and (3.18) becomes a linear one with respect to the function $q_n(x, t)$,

$$q_{\text{min}}(x, t) - \frac{q_{\text{min}}(x, t)}{2(q_{0-1, \text{min}}^{(n-1)}(x, 0))} + \frac{\partial q^{(0-1, \text{min}}_{n-1}(x, t) \partial_x q^{(0-1, \text{min}}_{n-1}(x, 0)}{2(q_{0-1, \text{min}}^{(n-1)}(x, 0))}^3 = 0 \text{ in } \Omega,$$  

$$q_\text{min}(\epsilon, t) = g_0(\epsilon + t), \quad q_{\text{min}}(\epsilon, t) = 2g_0(\epsilon + t), \quad q_{\text{min}}(b, t) = 0.$$  

To solve problem (4.6) and (4.7), we consider the following minimization problem.

**Minimization problem number $n$.** Assuming (3.12), minimize the functional $J^{(n)}_{\lambda, \beta} : H^2(\Omega) \to \mathbb{R}$ on the set $B(R, g_0)$,

$$J^{(n)}_{\lambda, \beta}(q_n) = \int_{\Omega} \left( q_{\text{min}}(x, t) - \frac{q_{\text{min}}(x, t)}{2(q_{0-1, \text{min}}^{(n-1)}(x, 0))} + \frac{\partial q^{(0-1, \text{min}}_{n-1}(x, t) \partial_x q^{(0-1, \text{min}}_{n-1}(x, 0)}{2(q_{0-1, \text{min}}^{(n-1)}(x, 0))}^3 \right)^2 e^{2\lambda \nu} \, dx \, dt + \beta \|q_n\|^2_{H^2(\Omega)}.$$  

Suppose that there exists unique minimizer $q_{0, \text{min}} \in \overline{B(R, g_0)}$ of the functional $J^{(n)}_{\lambda, \beta}(q_n)$. Then, following (3.19), (3.14) and (3.19), we set
The rest of the analytical part of this paper is devoted to the convergence analysis of the iterative numerical method presented in this section.

5. The Carleman estimate

In this section we formulate the Carleman estimate, which is the main tool of our construction. This estimate follows from theorem 3.1 of [24] as well as from (3.15) and (3.16). Let \( q(x,t) \in B(R,g_0) \) be an arbitrary function and let the function \( q_0(x,0) \) be constructed from the function \( q(x,t) \) as in (3.14). Consider the operator \( L_0 : H^2(\Omega) \rightarrow L^2(\Omega) \),

\[
L_0 u = u_{xx}(x,t) - u_{xt}(x,t) + \frac{1}{2} q_0^2(x,0),
\]

for all \( (x,t) \in \Omega \).

Theorem 5.1 (Carleman estimate [24]). There exists a number \( \alpha_0 = \alpha_0(R,\Omega) > 0 \) depending only on \( R,\Omega \) such that for any \( \alpha \in (0,\alpha_0) \) there exists a sufficiently large number \( \lambda_0 = \lambda_0(R,\Omega,\alpha) > 1 \) depending only on \( R,\Omega,\alpha \) such that for all \( \lambda \geq \lambda_0 \) and for all functions \( v \in H^2(\Omega) \) the following Carleman estimate holds:

\[
\int_{\Omega} (L_0 v)^2 e^{2\lambda_0 x} \, dx \, dt \\
\geq C \int_{\Omega} (\lambda (v_x^2 + v_t^2) + \lambda^3 v^2) e^{2\lambda_0 x} \, dx \, dt + C \int_{\varepsilon}^b (\lambda v_x^2(x,0) + \lambda^3 v^2(x,0)) e^{-2\lambda x} \, dx \\
- C \int_0^T (\lambda (v_x^2(\varepsilon,t) + v_t^2(\varepsilon,t)) + \lambda^3 v^2(\varepsilon,t)) e^{-2\lambda(\varepsilon+\alpha t)} \, dt \\
- C \int_{\varepsilon}^b (\lambda v_x^2(x,T) + \lambda^3 v^2(x,T)) e^{-2\lambda(x+\alpha T)} \, dx.
\]

Remark 5.1. Here and everywhere below \( C = C(R,\Omega,\alpha) > 0 \) denotes different constants depending only on listed parameters.

6. Strict convexity of functional (4.8) on the set \( B(R,g_0) \), existence and uniqueness of its minimizer

Functional (4.8) is quadratic. We prove in this section that it is strictly convex on the set \( B(R,g_0) \). In addition, we prove existence and uniqueness of its minimizer on this set. Although similar results were proven in many of the above cited publications on the convexification, see, e.g., [24] for the closest one, there are some peculiarities here, which are important for our convergence analysis, see remarks 6.1 and 6.2 below.

Introduce the subspace \( H^1_0(\Omega) \subset H^1(\Omega) \) as:

\[
H^1_0(\Omega) = \left\{ v \in H^1(\Omega) : v(\varepsilon,t) = v_4(\varepsilon,t) = v_4(b,t) = 0 \right\}.
\]
Denote $[,]$ the scalar product in the space $H^4(\Omega)$. Also, denote
\begin{equation}
A_n(q)(x, t) = q_{xx}(x, t) - q_{xx}(x, t) \frac{1}{2(q_{n-1, \min}^0(x, 0))}, 
\end{equation}
\begin{equation}
B_n(x, t) = \partial_t q_{(s-1), \min}(x, t) \frac{\partial_t q_{(s-1), \min}(x, 0)}{2(q_{n-1, \min}^0(x, 0))}.
\end{equation}

**Theorem 6.1.** Let $J_{\lambda, \beta}^{(0)}$ be the functional defined in (4.8). Then:

(a) For any set of parameters $\lambda, \beta$ and for any $q \in \overline{B(R, g_0)}$ this functional has the Frechet derivative $\left(J_{\lambda, \beta}^{(0)}(q)\right)' \in H_0^4(\Omega)$. The formula for $\left(J_{\lambda, \beta}^{(0)}(q)\right)'$ is:
\begin{equation}
\left(J_{\lambda, \beta}^{(0)}(q)\right)'(h) = 2 \int_\Omega \left( A_n(q)(x, t) + B_n(x, t) \right) A_n(h)(x, t) \frac{e^{2\lambda t}}{\lambda} \, dx \, dt + 2\beta \left[ q, h \right], \quad \forall \, h \in H_0^4(\Omega).
\end{equation}

(b) This derivative is Lipschitz continuous in $\overline{B(R, g_0)}$, i.e. there exists a constant $D > 0$ such that
\begin{equation}
\left\| \left(J_{\lambda, \beta}^{(0)}(q_2)\right)' - \left(J_{\lambda, \beta}^{(0)}(q_1)\right)' \right\|_{H_0^4(\Omega)} \leq D \left\| q_2 - q_1 \right\|_{H_0^4(\Omega)}, \quad \text{for all } q_1, q_2 \in \overline{B(R, g_0)}.
\end{equation}

(c) Let $\alpha_0 = \alpha_0(R, \Omega) > 0$, $\alpha \in (0, \alpha_0)$ and $\lambda_0 = \lambda_0(R, \Omega, \alpha) > 1$ be the numbers of theorem 5.1. Then there exists a sufficiently large constant
\begin{equation}
\lambda_1 = \lambda_1(R, \Omega, \alpha) \geq \lambda_0 > 1
\end{equation}
depending only on listed parameters such that for all $\lambda \geq \lambda_1$ and for all $\beta \in \left[2e^{-\lambda_0 t}, 1\right]$ the functional $J_{\lambda, \beta}^{(0)}(q)$ is strictly convex on the set $\overline{B(R, g)}$. More precisely, let $q \in \overline{B(R, g)}$ be an arbitrary function and also let the function $q + h \in \overline{B(R, g)}$. Then the following inequality holds:
\begin{equation}
J_{\lambda, \beta}^{(0)}(q + h) - J_{\lambda, \beta}^{(0)}(q) - \left(J_{\lambda, \beta}^{(0)}(q)\right)'(h) \geq C \int_\Omega \left[ \lambda \left( h_1^2 + h_2^2 \right) + \lambda^3 \lambda^2 \right] e^{2\lambda t} \, dx \, dt + C \int_\Omega \left( \lambda h_1^2(x, 0) + \lambda^3 h_2^2(x, 0) \right) e^{-2\lambda t} \, dx + \frac{\beta}{2} \left\| h \right\|_{H_0^4(\Omega)}^2, \quad \forall \, \lambda \geq \lambda_1.
\end{equation}

(d) For any $\lambda \geq \lambda_1$ there exists unique minimizer
\begin{equation}
q_{n, \min} \in \overline{B(R, g_0)}
\end{equation}
of the functional $J_{\lambda, \beta}^{(0)}(q)$ on the set $\overline{B(R, g_0)}$. Furthermore, the following inequality holds:
\begin{equation}
\left[ \left(J_{\lambda, \beta}^{(0)}(q)\right)'(q - q_{n, \min}) \right] \geq 0, \quad \forall \, q \in \overline{B(R, g_0)}.
\end{equation}
Remark 6.1. Even though the expression in the right-hand side of (6.4) is linear with respect to the function \( q \), we cannot use Riesz theorem here to prove existence and uniqueness of the minimizer \( q_{n, \text{min}} \), at which \( \left( J_{n, \text{min}}^{(0)} \right)' = 0 \). Rather, all what we can prove is (6.9). This is because we need to ensure that the function \( q_{n, \text{min}} \in \overline{B(R, g_0)} \).

Proof of Theorem 6.1. Since both function \( q, q + h \in \overline{B(R, g)} \) satisfy the same boundary conditions, then
\[
h \in H^2_0(\Omega). \tag{6.10}
\]
By (4.8) and (6.10)
\[
J_{n, \text{min}}^{(0)}(q + h) - J_{n, \text{min}}^{(0)}(q) = 2 \int_{\Omega} (A_n(q)(x, t) + B_n(x, t)) A_n(h)(x, t) e^{2\lambda \sigma} dx \, dt + 2\beta [q, h] \\
+ \int_{\Omega} [A_n(h)(x, t)]^2 e^{2\lambda \sigma} dx \, dt + \beta \| h \|^2_{H^1(\Omega)}, \quad \forall \ h \in H^2_0(\Omega). \tag{6.11}
\]
The expression in the first line of (6.11) coincides with expression in the right-hand side of (6.4). In fact, this is a bounded linear functional mapping \( H^2_0(\Omega) \) in \( \mathbb{R} \). Therefore, by Riesz theorem, there exists unique function \( \tilde{J}_n(q) \in H^2_0(\Omega) \) such that
\[
\tilde{J}_n(q), h = 2 \int_{\Omega} (A_n(q)(x, t) + B_n(x, t)) A_n(h)(x, t) e^{2\lambda \sigma} dx \, dt + 2\beta [q, h], \quad \forall \ h \in H^2_0(\Omega). \tag{6.12}
\]
In addition, it is clear from (6.11) and (6.12) that
\[
\lim_{\| h \|^2_{H^1(\Omega)} \to 0} \left\{ \frac{1}{\| h \|^2_{H^1(\Omega)}} \left[ J_{n, \text{min}}^{(0)}(q + h) - J_{n, \text{min}}^{(0)}(q) - \left[ \tilde{J}_n(q), h \right] \right] \right\} = 0.
\]
Therefore,
\[
\tilde{J}_n(q) = \left( J_{n, \text{min}}^{(0)} \right)'(q) \in H^2_0(\Omega) \tag{6.13}
\]
is the Fréchet derivative of the functional \( J_{n, \text{min}}^{(0)}(q) : \overline{B(R, g_0)} \to \mathbb{R} \) at the point \( q \in \overline{B(R, g_0)} \), and the right-hand side of (6.4) indeed represents \( \left( J_{n, \text{min}}^{(0)}(q) \right)'(h) \). Estimate (6.5) obviously follows from (6.4).

We now prove strict convexity estimate (6.7). To do this, we apply Carleman estimate (5.1) to the third line of (6.11). We obtain
\[
J_{n, \text{min}}^{(0)}(q + h) - J_{n, \text{min}}^{(0)}(q) - \left( J_{n, \text{min}}^{(0)}(q) \right)'(h) \\
\geq C \int_{\Omega} \left( \lambda (h^2_\epsilon + h^2_\sigma) + \lambda^2 h^2 \right) e^{2\lambda \sigma} dx \, dt + C \int_{\epsilon}^{0} \left( \lambda h^2_\epsilon(x, 0) \\
+ \lambda^2 h^2(x, 0) \right) e^{-2\lambda \sigma} dx - C \int_{\epsilon}^{0} \left( \lambda h^2_\epsilon(x, T) + \lambda^2 h^2(x, T) \right) e^{-2\lambda \sigma(x + \alpha T)} dx \\
+ \beta \| h \|^2_{H^1(\Omega)}, \quad \forall \ h \in H^2_0(\Omega). \tag{6.14}
\]
By trace theorem there exists a constant $C_1 = C_1(\Omega) > 0$ depending only on the domain $\Omega$ such that

$$\|v\|_{H^4(\Omega)}^2 \geq C_1 \|v(x, T)\|_{H^1(\Omega)}^2, \quad \forall \ v \in H^4(\Omega).$$

Since the regularization parameter $\beta \in \left[2e^{-\lambda\alpha T}, 1\right)$, then we can choose $\lambda_1$ so large that

$$\frac{\beta}{2} C_1 \geq C_1 e^{-\lambda\alpha T}, \quad \forall \ \lambda \geq \lambda_1, \ \forall \ x \in [\varepsilon, b].$$

Hence, for these values of $\lambda$, the expression in the last line of (6.14) can be estimated from the below as:

$$- C \int_{\varepsilon}^{b} \left( \lambda h_2^2(x, T) + \lambda^3 h_2^2(x, T) \right) e^{-2\lambda(x + \alpha T)} dx$$

$$+ \frac{\beta}{2} \|h\|_{H^4(\Omega)}^2 \geq \frac{\beta}{2} \|h\|_{H^4(\Omega)}^2, \quad \forall \ h \in H^4_0(\Omega). \quad (6.15)$$

Hence, (6.14) and (6.15) imply

$$J_{\lambda, \beta}^{(n)}(q + h) - J_{\lambda, \beta}^{(n)}(q) - \left(J_{\lambda, \beta}^{(n)}(q)\right)'(h) \geq C \int_{\Omega} \left( \lambda \left( h_2^2 + h_3^2 \right) + \lambda^3 h_2^2 \right) e^{2\lambda x} dx$$

$$+ C \int_{\varepsilon}^{b} \left( \lambda h_2^2(x, 0) + \lambda^3 h_2^2(x, 0) \right) e^{-2\lambda x} dx$$

$$+ \frac{\beta}{2} \|h\|_{H^4(\Omega)}^2, \quad \forall \ \lambda \geq \lambda_1.$$
7.1. Gradient descent method

Keeping in mind (3.12), choose an arbitrary function \( q_{0,n} \in B(R, g_0) \). We arrange the gradient descent method of the minimization of functional (4.8) as follows:

\[
q_{k,n} = q_{k-1,n} - \eta \left( J_n^{(0)}(q_{k-1,n}) \right), \quad k = 1, 2, \ldots,
\]

where \( \eta \in (0, 1) \) is a small number, which is chosen later. It is important to note that since functions \( J_n^{(0)}(q_{k-1,n}) \in H^1_0(\Omega) \), then boundary conditions (4.7) are kept the same for all functions \( q_{k,n}(x, t) \). Also, using (3.14) and (4.9), we set

\[
c_{k,n}(x) = \left( \frac{1}{2q_n^0(x, 0)} \right)^3, \quad x \in [\varepsilon, b],
\]

\[
c_{0,\text{min}}(x) = \left( \frac{1}{2q_{n,\text{min}}(x, 0)} \right)^3, \quad x \in [\varepsilon, b].
\]

Theorem 7.1 claims the global convergence of the gradient descent method (7.1) to the pair \( \{q_{n,\text{min}}, c_{n,\text{min}}\} \) in the case when \( q_{n,\text{min}} \in B(R/3, g_0) \), see remark 7.2.

**Theorem 7.1.** Let the number \( \lambda_1 = \lambda_1(R, \Omega, \alpha) \geq \lambda_0 > 1 \) be the one defined in (6.6). Let \( \lambda \geq \lambda_1 \). For this value of \( \lambda \), let \( q_{n,\text{min}} \in B(R, g_0) \) be the unique minimizer of the functional \( J_n^{(0)}(q_n) \) on the set \( B(R, g_0) \) with the regularization parameter \( \beta \in [2 \varepsilon^{-\lambda_0 T}, 1] \) (theorem 6.1). Assume that the function \( q_{n,\text{min}} \in B(R/3, g_0) \). For each \( n \), choose the starting point of the gradient descent method (7.1) as \( q_{0,n} \in B(R/3, g_0) \). Then, there exists a number \( \eta_n \in (0, 1) \) such that for any \( \eta \in (0, \eta_n) \) all functions \( q_{k,n} \in B(R, g_0) \). Furthermore, there exists a number \( \theta = \theta(\eta) \in (0, 1) \) such that the following convergence estimates are valid:

\[
\|q_{k,n} - q_{n,\text{min}}\|_{H^1(\Omega)} \leq \theta^k \|q_{0,n} - q_{n,\text{min}}\|_{H^1(\Omega)}, \quad k = 1, \ldots,
\]

\[
\|c_{k,n} - c_{n,\text{min}}\|_{H^1(\Omega)} \leq C\theta^k \|q_{0,n} - q_{n,\text{min}}\|_{H^1(\Omega)}, \quad k = 1, \ldots,
\]

**Proof.** Estimate (7.4) follows immediately from [26, theorem 4.6] combined with ‘corrections’ of functions \( q_{k,n}(x, 0) \), \( q_{n,\text{min}}(x, 0) \) via (3.14). Estimate (7.5) follows immediately from trace theorem, (3.14) and (7.2)–(7.4).

7.2. Gradient projection method

Suppose now that there is no information on whether or not the function \( q_{n,\text{min}} \in B(R/3, g_0) \). In this case we construct the gradient projection method. We introduce the function \( F(x, t) \) below since it is easy to construct the projection operator on a ball with the center at \( \{0\} \).

Consider the function \( \chi(x) \) such that

\[
\chi(x) \in C^4[\varepsilon, b], \quad \chi(x) = \begin{cases} 1, & x \in [\varepsilon, b/4], \\ 0, & x \in [b/2, b], \\ \text{between 0 and 1} & \text{for } x \in (b/4, b/2). \end{cases}
\]

The existence of such functions \( \chi(x) \) is well known from the real analysis course. Suppose that the function \( g_0(t) \in H^2(0, T + \varepsilon) \). Define the function \( F \in H^1(\Omega) \) as

\[
F(x, t) = \chi(x) \left( g_0(t + \varepsilon) + 2xg_0'(t + \varepsilon) \right).
\]
Then
\[
F(\varepsilon, t) = g_0(t + \varepsilon), \quad F_x(\varepsilon, t) = 2g'_0(t + \varepsilon), \quad F_x(b, t) = 0. 
\]  
(7.6)
Denote
\[
p_n(x, t) = q_n(x, t) - F(x, t). 
\]  
(7.7)
Then (4.6) and (4.7) become:
\[
p_{nxx}(x, t) = \frac{1}{2} \frac{1}{(q_{n-1, \min}(x, 0))^2} + F_{xx}(x, t) - F_x(x, t) + \frac{1}{2} \frac{1}{(q_{n-1, \min}(x, 0))^2} + \partial_x q_{n-1}(x, t) \frac{\partial_x q_{n-1}(x, 0)}{2(q_{n-1, \min}(x, 0))} = 0 \quad \text{in } \Omega, 
\]  
(7.8)
(7.9)
Assume that
\[
\|F\|_{H^1(\Omega)} < R. 
\]  
(7.10)
By (7.6), (7.7), (7.9) and (7.10) and triangle inequality
\[
p_n \in B_0(2R) = \left\{ p \in H^1_0(\Omega) : \|p\|_{H^1_0(\Omega)} < 2R \right\}. 
\]  
(7.11)
To find the function \( p_n \in B_0(2R) \) satisfying conditions (7.8) and (7.9), we minimize the following functional \( I_{\lambda, \beta}^{(n)}(p_n) : B_0(2R) \to \mathbb{R} \)
\[
I_{\lambda, \beta}^{(n)}(p_n) = I^{(n)}_{\lambda, \beta}(p_n + F), \quad p_n \in B_0(2R). 
\]  
(7.12)
**Remark 7.1.** An obvious analog of theorem 6.1 is valid of course for the functional \( I_{\lambda, \beta}^{(n)}(p_n) \) defined in (7.12). But in this case we should have instead of (6.6) \( \lambda \geq \lambda_1 = \lambda_1(2R, \Omega, \alpha) \geq \lambda_0 > 1. \) In particular, there exists unique minimizer \( p_{n, \min} \in B_0(2R) \) of this functional on the closed ball \( B_0(2R) \). We omit the formulation of this theorem, since it is an obvious reformulation of theorem 6.1.

Let \( P_{B_0(2R)} : H^1_0(\Omega) \to B_0(2R) \) be the projection operator of the space \( H^1_0(\Omega) \) on the closed ball \( B_0(2R) \subset H^1_0(\Omega) \). Then this operator can be easily constructed:
\[
P_{B_0(2R)}(p) = \begin{cases} 
p & \text{if } p \in B_0(2R), \\
\frac{2R}{\|p\|_{H^1_0(\Omega)}}p & \text{if } p \notin B_0(2R). 
\end{cases} 
\]
We now construct the gradient projection method of the minimization of the functional \( I_{\lambda, \beta}^{(n)}(p_n) \) on the set \( B_0(2R) \). Let \( p_{0, n} \in B_0(2R) \) be an arbitrary function. Then the sequence of the gradient projection method is:
\[
p_{k, n} = P_{B_0(2R)} \left( p_{(k-1), n} - \eta \left( I_{\lambda, \beta}^{(n)}(p_{(k-1), n}) \right) \right), \quad k = 1, 2, \ldots, 
\]  
(7.13)
where \( \eta \in (0, 1) \) is a small number, which is chosen later. Using (7.2), (7.3) and (7.7), we set
\[
\tilde{c}_{k,n}(x) = \frac{1}{(2q^0_{h,k}(x,0))^4}, \quad x \in [\varepsilon, b], \tag{7.14}
\]
\[
\tilde{c}_{n,\text{min}}(x) = \frac{1}{(2q^0_{n,\text{min}}(x,0))^4}, \quad x \in [\varepsilon, b]. \tag{7.15}
\]
Here the function \(q^0_{h,k}(x,0)\) is obtained as follows: first, we consider the function \((p_{h,k} + F)(x,0)\). Next, we apply procedure (3.14) to this function. Similarly for \(q^0_{n,\text{min}}(x,0)\).

Denote \(p_{n,\text{min}} \in B_0(2R)\) the unique minimizer of functional (7.12) on the set \(B_0(2R)\) (remark 7.1). Following (7.7), denote
\[
\tilde{q}_{n,\text{min}} = p_{n,\text{min}} + F, \quad \tilde{q}_{k,n} = p_{k,n} + F. \tag{7.16}
\]
We omit the proof of theorem 7.2 since it is very similar with the proof of theorem 7.1. The only difference is that instead of theorem 4.6 of [26] one should use theorem 4.1 of [24].

**Theorem 7.2.** Let (7.6) and (7.10) hold. Let the number \(\lambda_1 = \lambda_1(R, \Omega, \alpha) \geq \lambda_0 > 1\) be the one defined in (6.6) and let
\[
\lambda \geq \bar{\lambda}_1 = 1 (2R, \Omega, \alpha) \geq \lambda_1 (R, \Omega, \alpha).
\]

For this value of \(\lambda\) and for \(\beta \in \left[2 e^{-\lambda T}, 1\right]\) let \(p_{n,\text{min}} \in B_0(2R)\) be the unique minimizer of functional (7.12) on the set \(B_0(2R)\) (remark 7.1). Let notations (7.14) hold. For each \(n\), choose the starting point of the gradient projection method (7.1) as \(p_{0,n} \in B_0(2R)\). Then there exists a number \(\eta_0 \in (0, 1)\) such that for any \(\eta \in (0, \eta_0)\) there exists a number \(\theta = \theta(\eta) \in (0, 1)\) such that the following convergence estimates are valid for the iterative process (7.13):
\[
\|\tilde{q}_{k,n} - \tilde{q}_{n,\text{min}}\|_{H^1(\Omega)} \leq \theta^k \|\tilde{q}_{0,n} - \tilde{q}_{n,\text{min}}\|_{H^1(\Omega)},
\]
\[
\|\tilde{c}_{k,n} - \tilde{c}_{n,\text{min}}\|_{H^1(\Omega)} \leq C\theta^k \|\tilde{q}_{0,n} - \tilde{q}_{n,\text{min}}\|_{H^1(\Omega)},
\]
where functions \(\tilde{c}_{k,n}, \tilde{c}_{n,\text{min}}, \tilde{q}_{k,n}, \tilde{q}_{n,\text{min}}\) are defined in (7.14)–(7.16).

**Remark 7.2.** Both theorems 7.1 and 7.2 claim the global convergence of corresponding versions of the gradient method to pairs \((q_{n,\text{min}}, c_{n,\text{min}})\) and \((\tilde{q}_{n,\text{min}}, \tilde{c}_{n,\text{min}})\). This is because the starting function in both cases is an arbitrary one either in \(B(R/3, g_0)\) or in \(B_0(2R)\) and a smallness condition is not imposed on the number \(R\). Also, see the second item of remark 8.1 and our definition of the global convergence in introduction.

### 8. Contraction mapping and global convergence

In this section, we prove the global convergence of the numerical method of section 4 for solving problem 3.1. To do this, we introduce first the exact solution of our CIP. Recall that the concept of the existence of the exact solution is one of the main concepts of the theory of ill-posed problems [5, 41]. In particular, an estimate in our global convergence theorem is very similar to the one in contraction mapping.

Suppose that there exists a function \(c^*(x)\) satisfying conditions (2.1) and (2.2). Let \(u^*(x, t)\) be the solution of problem (2.3) and (2.4) with \(c := c^*\). We assume that the corresponding
data \( g_0^* (t), \partial_1 g_0^* (t) \) for the CIP are noiseless, see (2.11) and (2.12). Let \( q^* (x, t) \) be the function \( q^* (x, t) \) which is constructed from the function \( u^* (x, t) \) as in (3.1). Following (3.11), we assume that

\[
q^* \in B(R, g_0^*),
\]

\[
B(R, g_0^*) = \left\{ q \in H^4(\Omega) : \|q\|_{H^4(\Omega)} < R, \begin{aligned}
q(\varepsilon, t + \varepsilon) &= g_0^*(t + \varepsilon), q_1(\varepsilon, t) = 2(g_0^*)'(t + \varepsilon), \\
q_4(b, t) &= 0, \\
\frac{1}{2b^{1/4}} &\leq q(x, 0) \leq 1/2, \ x \in [\varepsilon, b]
\end{aligned} \right\}. \tag{8.2}
\]

By (3.17) and (3.18)

\[
q^*_{xx}(x, t) - q^*_x(x, t) \frac{1}{2(q^*(x, 0))^2} + q^*_x(x, t) - q^*_t(x, 0) - q^*_t(x, 0) = 0 \quad \text{in} \ \Omega, \tag{8.3}
\]

\[
q^*(\varepsilon, t) = g_0^*(t + \varepsilon), \quad q^*_x(\varepsilon, t) = 2\partial_1 g_0^*(t + \varepsilon), \quad q^*_x(b, t) = 0. \tag{8.4}
\]

By (3.3)

\[
c^* (x) = \frac{1}{(2q^*(x, 0))^2}. \tag{8.5}
\]

It is important in the formulation of theorem 6.1 that both functions \( q \) and \( q + h \) should have the same boundary conditions as prescribed in \( B(R, g_0) \). However, boundary conditions for functions \( q_n \) and \( q^* \) are different. Hence, similarly to subsection 7.2, we consider a function \( F^* \in H^4(\Omega) \) such that (see (7.6))

\[
F^*(\varepsilon, t) = g_0^*(t + \varepsilon), \quad F^*_x(\varepsilon, t) = 2\partial_1 g_0^*(t + \varepsilon), \quad F^*_x(b, t) = 0. \tag{8.6}
\]

We assume similarly to (7.10) that

\[
\|F^*\|_{H^4(\Omega)} < R. \tag{8.7}
\]

Also, similarly to (7.7), we introduce the function \( p^* (x, t) \) as:

\[
p^*(x, t) = q^*(x, t) - F^*(x, t). \tag{8.8}
\]

Let \( \lambda_1 \) be the number defined in (6.6), let \( \lambda \geq \lambda_1 \) and the function \( q_{n,\min} \in B(R, g_0) \) (see (6.8)) be the unique minimizer of the functional \( J_{\lambda n}^* (q_n) \) on the set \( B(R, g_0) \), the existence of which is established in theorem 6.1. Following (7.7), denote

\[
p_{n,\min} (x, t) = q_{n,\min} (x, t) - F(x, t). \tag{8.9}
\]

By (6.8), (7.11), (8.1), (8.2) and (8.7)–(8.9)

\[
p_{n,\min, p^*} \in \overline{B_0 (2R)}. \tag{8.10}
\]

Also, it follows from embedding theorem, (3.13), (6.8), (7.10), (8.7) and (8.10) that

\[
\|q^*\|_{C^2(\overline{\Omega})}, \|p^*\|_{C^2(\overline{\Omega})}, \|q_{n,\min}\|_{C^2(\overline{\Omega})}, \|p_{n,\min}\|_{C^2(\overline{\Omega})} \leq C. \tag{8.11}
\]
We assume that the data $g_0, g'_0$ for our CIP are given with noise of the level $\delta$, where the number $\delta > 0$ is sufficiently small. More precisely, we assume that

$$\|F - F^*\|_{H^4(\Omega)} < \delta. \quad (8.12)$$

Observe that (3.14), (8.1) and (8.2) imply that

$$|q_{n-1, \text{min}}(x, 0) - q^* (x, 0)| \leq |q_{n-1, \text{min}}(x, 0) - q^* (x, 0)|, \quad x \in [\varepsilon, b]. \quad (8.13)$$

By (8.3), (8.4), (8.6) and (8.8)

$$p^*_\varepsilon(x, t) - p^*_\varepsilon(x, t) + q^*_\varepsilon(x, t) \frac{q^*_\varepsilon(x, 0)}{2(q^*_\varepsilon(x, 0))^2} + F^*_{xt} - F^*_\varepsilon(x, t) \frac{1}{2(q^*_\varepsilon(x, 0))^2} = 0 \quad (8.14)$$

in $\Omega \times [0, T]$ and

$$p^*(\varepsilon, t) = p^*_\varepsilon(x, t) = p^*_\varepsilon(b, t) = 0. \quad (8.15)$$

**Theorem 8.1 (Contraction mapping and the global convergence of the method of section 3).** Let functions $F, F^* \in H^4(\Omega)$ satisfy conditions (7.6), (7.10), (8.6), (8.7) and (8.12). Let a sufficiently large number $\lambda_1 = \lambda_1(R, \Omega, \alpha) \geq \lambda_0 > 1$ be the one defined in (6.6). Let

$$\lambda \geq \tilde{\lambda}_1 = \lambda_1(2R, \Omega, \alpha) \geq \lambda_1(R, \Omega, \alpha). \quad (8.16)$$

For this value of $\lambda$, let $q_{n, \text{min}} \in B(R, \Omega_0)$ be the unique minimizer of the functional $J^{(n)}_{\lambda}(q_n)$ on the set $B(R, \Omega_0)$ with the regularization parameter $\beta \in [2e^{-\lambda_0T}, 1]$ (Theorem 6.1). Let

$$\overline{q}_n = q_{n, \text{min}} - q^*, \quad \overline{q}_n = c_{n, \text{min}} - c^*, \quad (8.17)$$

where $c_{n, \text{min}}$ is defined in (7.3). Then the following convergence estimate holds

$$\int_{\Omega} (\overline{q}^2_{nx} + \overline{q}^2_{nt} + \overline{q}^2_n) (x, t) e^{2\lambda \varepsilon} \, dx \, dt + \int_0^b (\overline{q}^2_{nx} + \overline{q}^2_n) (x, 0) e^{-2\lambda \varepsilon} \, dx \leq \frac{C}{\lambda} \int_\Omega (\overline{q}_{(n-1)x}^2 + \overline{q}_{(n-1)t}^2 + \overline{q}_{n-1}) (x, t) e^{2\lambda \varepsilon} \, dx \, dt \quad (8.18)$$

which leads to:

$$\int_{\Omega} (\overline{q}^2_{nx} + \overline{q}^2_{nt} + \overline{q}^2_n) (x, t) e^{2\lambda \varepsilon} \, dx \, dt \leq \frac{C^n}{\lambda^4} \int_{\Omega} (\overline{q}_{nx}^2 + \overline{q}_{nt}^2 + \overline{q}_n^2) (x, t) e^{2\lambda \varepsilon} \, dx \, dt + C (\delta^2 + \beta). \quad (8.19)$$

In addition,

$$\|\overline{q}_n\|_{H^4(\varepsilon,b)}^2 \leq \frac{C^n}{\lambda^4} \int_{\Omega} (\overline{q}_{nx}^2 + \overline{q}_{nt}^2 + \overline{q}_n^2) (x, t) e^{2\lambda \varepsilon} \, dx \, dt + C (\delta^2 + \beta). \quad (8.20)$$
Remark 8.1.

(a) Due to the presence of the term $C/\lambda$ with a sufficiently large $\lambda$, estimate (8.18) is similar to the one in the classical contraction mapping principle, although we do not claim here the existence of the fixed point.

(b) When computing the unique minimizer $q_{\min}$ of functional (4.4) on the set $\mathcal{B}(K, \Omega)$, we do not impose a smallness condition on the number $R$. Therefore, theorem 8.1 claims the global convergence of the method of section 4: see our definition of the global convergence in introduction. The same is true for theorems 9.1 and 9.2 in section 9.

**Proof Theorem 8.1.** Denote $h_n = p^* - p_{n, \min}$. By (8.9) $h_n = -q_n + (F - F^*)$. Hence, (8.12) and embedding theorem imply:

$$h_n^2 \geq \frac{1}{2} \mathbf{\tau}_n^2 - C \delta^2, h_{nx}^2 \geq \frac{1}{2} \mathbf{\tau}_{nx}^2 - C \delta^2, h_{nt}^2 \geq \frac{1}{2} \mathbf{\tau}_{nt}^2 - C \delta^2 \quad \text{in } \Omega_{\delta}$$  \hspace{1cm} (8.21)

$$h_n^2 + h_{nx}^2 + h_{nt}^2 \leq C (\mathbf{\tau}_n^2 + \mathbf{\tau}_{nx}^2 + \mathbf{\tau}_{nt}^2 + \delta^2) \quad \text{in } \Omega_{\delta}.$$  \hspace{1cm} (8.22)

Consider the functional $I_{\lambda, \beta}^{(n)}(p^*) - I_{\lambda, \beta}(p_{n, \min}) - \left(I_{\lambda, \beta}(p_{n, \min})\right)'(h_n)$

$$\geq C \int \left[ \lambda (h_{nx}^2 + h_{nt}^2) + \lambda^3 h_{nt}^2 \right] e^{2 \lambda \|h_n\|_{L^2(\Omega)}} dt$$

$$+ C \int_{\Omega} \left( \lambda h_{nt}^2(x, 0) + \lambda^3 h_{nt}^2(x, 0) \right) e^{-2 \lambda x} dx + \frac{\beta}{2} \|h_n\|_{L^2(\Omega)}, \quad \forall \lambda \geq \lambda_1.$$  \hspace{1cm} (8.23)

By (6.9)

$$- \left(I_{\lambda, \beta}(p_{n, \min})\right)'(h_n) \leq 0.$$  \hspace{1cm} (8.24)

Hence, the left-hand side of (8.23) can be estimated as:

$$I_{\lambda, \beta}^{(n)}(p^*) - I_{\lambda, \beta}(p_{n, \min}) - \left(I_{\lambda, \beta}(p_{n, \min})\right)'(h_n) \leq I_{\lambda, \beta}^{(n)}(p^*).$$  \hspace{1cm} (8.24)

We now estimate the right-hand side of (8.24) from the above. It follows from (4.8), (7.12) and (8.3) that

$$I_{\lambda, \beta}^{(n)}(p^*) = \int_{\Omega} G_n e^{2 \lambda x} dx + \beta \|p^* - F\|_{L^2(\Omega)},$$  \hspace{1cm} (8.25)

where

$$G_n = p_{nx}^{(n)}(x, t) - p_{nx}^{(n)}(x, t) - \frac{1}{2(q_{(0)-1,\min}(x, 0))^2} \frac{\partial \delta \varphi_{(n)-1,\min}(x, t)}{2(q_{(0)-1,\min}(x, 0)^2)}$$

$$+ F_{xx} - F_{xx}(x, t) - \frac{1}{2(q_{(0)-1,\min}(x, 0)^2)}$$

$$= q_{xx}^{(n)}(x, t) - q_{xx}^{(n)}(x, t) - \frac{1}{2(q^*(x, 0))^2} + q_{xx}^{(n)}(x, t) - \frac{q_{xx}^*(x, 0)}{2(q^*(x, 0))^2}.$$  \hspace{1cm} (8.26)
Using (8.12) and (8.13), we obtain

\[ + (F - F^*)_{xx} - (F - F^*)_{tt} + \frac{1}{2} \frac{1}{(q(x, 0))^2} \]

\[ - q_{tt} \left( \frac{1}{2} \frac{1}{(q(x, 0))^2} - \frac{1}{2} \frac{1}{(q^*(x, 0))^2} \right) \]

\[ + \left( \frac{\partial \tilde{q}(t, t)}{2(q(x, 0))^2} - q_{tt} \right) \frac{q^*(x, 0)}{2(q^*(x, 0))^2} \] .

By (8.3), the third line of (8.26) equals zero. Hence, (8.26) becomes

\[ G_n = (F - F^*)_{xx} + (F - F^*)_{tt} + \frac{1}{2} \frac{1}{(q(x, 0))^2} \]

\[ - q_{tt} \left( \frac{1}{2} \frac{1}{(q(x, 0))^2} - \frac{1}{2} \frac{1}{(q^*(x, 0))^2} \right) \]

\[ + \left( \frac{\partial \tilde{q}(t, t)}{2(q(x, 0))^2} - q_{tt} \right) \frac{q^*(x, 0)}{2(q^*(x, 0))^2} \] .

Hence, by (3.14) and (8.12) and embedding theorem

\[ |G_n| \leq C \delta + C \left| \frac{1}{2(q(x, 0))^2} - \frac{1}{2(q^*(x, 0))^2} \right| \]

\[ + \left| \frac{\partial \tilde{q}(t, t)}{2(q(x, 0))^2} - q_{tt} \right| \frac{q^*(x, 0)}{2(q^*(x, 0))^2} \] . \hspace{1cm} (8.27)

Using (8.12) and (8.13), we obtain

\[ \frac{1}{2(q(x, 0))^2} - \frac{1}{2(q^*(x, 0))^2} \]

\[ = \left| \frac{q(x, 0) - q^*(x, 0)}{2(q(x, 0))^2} - (F - F^*) (x, 0) \right| \frac{q(x, 0) + q^*(x, 0)}{2(q(x, 0))^2} \]

\[ \leq C \delta + C |h_{n-1}(x, 0)| . \]

Combining this with (8.27), we obtain

\[ |G_n| \leq C \delta + C |h_{n-1}(x, 0)| \]

\[ + \left| \frac{\partial \tilde{q}(t, t)}{2(q(x, 0))^2} - q_{tt} \right| \frac{q^*(x, 0)}{2(q^*(x, 0))^2} \] . \hspace{1cm} (8.28)

Next,
\[
\frac{1}{2(q_{(n-1)\min}^0(x, 0))^3} = \frac{1}{2(q^*(x, 0))^3} + \left(\frac{1}{2(q_{(n-1)\min}^0(x, 0))^3} - \frac{1}{2(q^*(x, 0))^3}\right)
\]

\[
= \frac{1}{2(q^*(x, 0))^3} + S_{n-1}(x, 0) \left[ (q_{(n-1)\min}^0(x, 0) - q^*(x, 0)) - (F - F^*) (x, 0) \right],
\]

where the function \( S_{n-1}(x, 0) \) can be estimated as

\[
|S_{n-1}(x, 0)| \leq C. \tag{8.29}
\]

Hence,

\[
\partial_t q_{(n-1)\min}^0(x, t) \frac{\partial_x q_{(n-1)\min}^0(x, 0)}{2(q_{(n-1)\min}^0(x, 0))^3} = \partial_t q_{(n-1)\min}^0(x, t) \frac{\partial_x q_{(n-1)\min}^0(x, 0)}{2(q^*(x, 0))^3} + \left[ \partial_t q_{(n-1)\min}^0(x, t) \partial_x q_{(n-1)\min}^0(x, 0) \right]
\]

\[
\times S_{n-1}(x, 0) \left[ (q_{(n-1)\min}^0(x, 0) - q^*(x, 0)) - (F - F^*) (x, 0) \right].
\]

Hence, using (3.11), (3.14), (8.1), (8.13) and (8.29), we obtain

\[
\left| \partial_t q_{(n-1)\min}^0(x, t) \frac{\partial_x q_{(n-1)\min}^0(x, 0)}{2(q_{(n-1)\min}^0(x, 0))^3} - q_t^*(x, 0) \frac{q_t^*(x, 0)}{2(q^*(x, 0))^3} \right|
\]

\[
\leq \frac{1}{2(q^*(x, 0))^3} |\partial_t q_{(n-1)\min}^0(x, t) \partial_x q_{(n-1)\min}^0(x, 0) - q_t^*(x, 0) q_t^*(x, 0)| \tag{8.30}
\]

\[
+ C \left( \delta + |h_{n-1}(x, 0)| \right)
\]

\[
\leq C \left( \delta + |h_{n-1}(x, 0)| + |\partial_x h_{n-1,x}(x, 0)| + |h_{n-1,x}(x, t)| \right).
\]

Combining (8.28) with (8.30), we obtain

\[
|G_n(x, t)| \leq C \left( \delta + |h_{n-1}(x, 0)| + |h_{n-1,x}(x, 0)| + |h_{n-1,x}(x, t)| \right).
\]

Hence, by (8.25)

\[
F_{\lambda, p}^n(p) \leq C \int_\Omega \left( \delta^2 + h_{n-1}^2(x, 0) + h_{n-1,x}^2(x, 0) + h_{n-1,t}^2(x, t) \right) e^{2\lambda_\beta} dx dt + C\beta.
\]

Substituting this in (8.24) and then using (8.23), we obtain

\[
\int_\Omega (h_{n,t}^2 + h_{n,x}^2 + h_{n,x}^2) e^{2\lambda_\beta} dx dt + \int_\Sigma (h_{n,t}^2(x, 0) + h_{n,x}^2(x, 0)) e^{-2\lambda_\beta} dx
\]

\[
\leq C \lambda \int_\Omega (\delta^2 + h_{n-1}^2(x, 0) + h_{n-1,x}^2(x, 0) + h_{n-1,t}^2(x, t)) e^{2\lambda_\beta} dx dt + C\beta. \tag{8.31}
\]
Obviously

\[ \int_{\Omega} \left( h_{n-1}^2(x,0) + h_{n-1}^2(x,0) \right) e^{2\lambda \varphi} \, dx \, dt \leq \frac{1}{2 \lambda \alpha} \int_{\varepsilon}^{b} \left( h_{0}^2(x,0) + h_{n}^2(x,0) \right) e^{-2\lambda \varphi} \, dx, \]

(8.32)

\[ \int_{\Omega} V^2 e^{2\lambda \varphi} \, dx \, dt \leq C \delta^2 e^{2\lambda \varphi}. \]

(8.33)

Denote

\[ y_n = \int_{\Omega} \left( h_{n}^2(x,0) + h_{n}^2(x,0) \right) e^{2\lambda \varphi} \, dx \, dt + \int_{\varepsilon}^{b} \left( h_{n+1}^2(x,0) + h_{n+1}^2(x,0) \right) e^{-2\lambda \varphi} \, dx. \]

(8.34)

Then (8.31)–(8.33) imply

\[ y_n \leq C \lambda y_{n-1}^{\lambda} + C \left( \frac{\delta^2}{\lambda^2} + \beta \right). \]

(8.35)

Iterating (8.35) with respect to \( n \), we obtain

\[ \int_{\Omega} \left( h_{n}^2(x,0) + h_{n}^2(x,0) \right) e^{2\lambda \varphi} \, dx \, dt + \int_{\varepsilon}^{b} \left( h_{n+1}^2(x,0) + h_{n+1}^2(x,0) \right) e^{-2\lambda \varphi} \, dx \]

\[ \leq C \lambda^2 \left[ \int_{\Omega} \left( h_{0}^2(x,0) + h_{0}^2(x,0) \right) e^{2\lambda \varphi} \, dx \, dt + \int_{\varepsilon}^{b} \left( h_{n}^2(x,0) + h_{n}^2(x,0) \right) e^{-2\lambda \varphi} \, dx \right] + C \left( \frac{\delta^2}{\lambda^2} + \beta \right). \]

(8.36)

Apply (8.21) to the left-hand side of estimate (8.36). Also, apply (8.22) at \( n = 0 \) to the right-hand side of (8.36). We obtain (8.19). Estimate (8.20) follows from an obvious combination of (8.19) with (7.3), (8.5), (8.13) and (8.17). Finally, estimate (8.18) follows immediately from (8.21), (8.22) and (8.31). □

9. Global convergence of the gradient and gradient projection methods to the exact solution

First, we consider the gradient method of the minimization of functionals \( J^{\lambda}(q_{k-1}, n) \) on the set \( B(R, g_0) \), see (7.1). The proof of theorem 9.1 follows immediately from the triangle inequality combined with theorems 7.1 and 8.1.

**Theorem 9.1.** Let \( \alpha_0 \) and \( \lambda_0 \) be the numbers of theorem 5.1. Let the sufficiently large number \( \lambda_1 = \lambda_1(R, \Omega, \alpha) \geq \lambda_0 > 1 \) be the one defined in (6.6). Let the number \( \tilde{\lambda}_1 \) be the same as in (8.16).

\[ \tilde{\lambda}_1 = \lambda_1(2R, \Omega, \alpha) \geq \lambda_1(R, \Omega, \alpha). \]

Let \( \lambda \geq \lambda_1 \) and let the regularization parameter \( \beta \in \left[ 2 e^{-\lambda \alpha T}, 1 \right] \). Assume that the functions \( q_{\min, n} \in B(R/3, g_0) \) for all \( n \). For each \( n \), choose the starting point of the gradient method (7.1) as \( q_{0,n} \in B(R/3, g_0) \). Then there exists a number \( \eta_0 \in (0, 1) \) such that for any \( \eta \in (0, \eta_0) \)
functions $q_{n,k} \in B(R, g_0), \forall k, n = 1, \ldots$. Furthermore, there exists a number $\theta = \theta(\eta) \in (0, 1)$ such that the following convergence estimate is valid:

\[
\|c_{k,n} - c^*\|_{H^{1/2}(\tilde{\Omega})} \leq C 0^n \|q_{0,n} - q_{n,\min}\|_{H^1(\Omega)} + C n/2 \left( \int_\Omega (\overline{q_{0,n}}^2 + \overline{q_0}^2 + \overline{q_0}^2) \cdot (x, t) e^{2b\sqrt{\beta} |x|} \, dx \, dt \right)^{1/2} + C \left( \delta + \sqrt{\beta} \right),
\]

where the function $\tilde{\eta}_0$ is defined in (8.17).

Consider now the gradient projection method of the minimization of the functionals $I_{\lambda}(p_n) = I_{\lambda}(p_n + F)$ in (7.12) on the set $B_0(2R)$. We use notations (7.14). Theorem 9.2 follows immediately from the triangle inequality combined with theorems 7.2 and 8.1.

**Theorem 9.2.** Let the number $\lambda_1 = \lambda_1(R, \Omega, x) \geq 1$ be the one defined in (6.6). Let the number $\chi_1$ be the same as in (8.16),

\[
\pi_1 = \lambda_1(2R, \Omega, x) \geq \lambda_1(R, \Omega, x).
\]

Let $\lambda \geq \lambda_1$ and let the regularization parameter $\beta \in [2e^{-\lambda_0 T}, 1]$. Consider the gradient projection method (7.13). For each $n$, choose the starting point $p_{0,n}$ of this method as an arbitrary point of the ball $B_0(2R)$. Then there exists a number $\eta_0 \in (0, 1)$ such that for any $\eta \in (0, \eta_0)$ there exists a number $\theta = \theta(\eta) \in (0, 1)$ such that the following convergence estimate is valid:

\[
\|\tilde{c}_{k,n} - c^*\|_{H^{1/2}(\tilde{\Omega})} \leq C 0^n \|\tilde{q}_{0,n} - \tilde{q}_{\min}\|_{H^1(\Omega)} + C n/2 \left( \int_\Omega (\overline{\tilde{q}_{0,n}}^2 + \overline{\tilde{q}_0}^2 + \overline{\tilde{q}_0}^2) \cdot (x, t) e^{2b\sqrt{\beta} |x|} \, dx \, dt \right)^{1/2} + C \left( \delta + \sqrt{\beta} \right),
\]

where functions $\tilde{c}_{k,n}, \tilde{q}_{0,n}, \tilde{q}_{\min}$ and $\tilde{\eta}_0$ are defined in (7.14), (7.16) and (8.17) respectively.

### 10. Numerical studies

#### 10.1. Numerical implementation

To generate the simulated data, we use lemma 2.2. This means that we solve problem (2.3) and (2.4) for the case when the whole real line is replaced by a large interval $(-a, a)$ with the absorbing boundary conditions (2.9) and (2.10). More precisely, just as in section 6.1 of [24], we use the implicit scheme to numerically solve

\[
\begin{cases}
\begin{aligned}
c(x)u_n(x, t) &= u_{xx}(x, t), \\
u(-a, t) &= u_{xx}(-a, t) = 0 & t \in (0, T), \\
u(a, t) &+ u_{xx}(a, t) = 0 & t \in (0, T), \\
u(x, 0) &= 0 & x \in \mathbb{R}, \\
u_t(x, 0) &= \tilde{\delta}(x) & x \in \mathbb{R},
\end{aligned}
\end{cases}
\tag{10.1}
\]

where $a = 5$, $T = 6$ and

\[
\tilde{\delta}(x) = \frac{30}{\sqrt{2\pi}} e^{-\tau_0 x^2}
\]
Algorithm 1. A numerical method to solve problem 3.1.

1: Choose a set of parameters $\lambda$, $\alpha$ and $\beta$
2: Compute the function $q_0$ by minimizing the functional $J_{\lambda,\beta}^{(0)}$ defined in (4.4). Due to (3.3), the initial reconstruction is given by
   \[ c_{\text{init}}(x) = \frac{1}{2q_0(x)\Delta x^2} \quad \text{for all } x \in [\epsilon, b] \]
3: Assume that the function $q_{n-1}$ is known. We compute the function $q_n$ by minimizing the functional $J_{\lambda,\beta}^{(n)}$ defined in (4.8)
4: Set $q_{\text{comp}} = q_n$ when $n = n^*$ is large enough
5: Due to (3.3), the function $c_{\text{comp}}$ is set to be
   \[ c_{\text{comp}}(x) = \frac{1}{2q_{\text{comp}}(x)\Delta x^2} \quad \text{for all } x \in [\epsilon, b] \]

is a smooth approximation of the function $\delta(x)$. We solve problem (10.1) by the implicit finite difference method. In the finite difference scheme, we arrange a uniform partition for the interval $[-a, a]$ by $\{y_i = -a, y_1, \ldots, y_N = a\} \subset [-a, a]$ with $y_i = a + 2i\Delta a/N_t$, $i = 0, \ldots, N_t$, where $N_t$ is a large number. In the time domain, we split the interval $[0, T]$ into $N_t + 1$ uniform sub-intervals $[t_j, t_{j+1}]$, $j = 0, \ldots, N_t$, with $t_j = jT/N_t$, where $N_t$ is a large number. In our computational setting, $N_t = 3001$ and $N_i = 301$. These numbers are the same as in [24].

We observe a computational error for the function $u$ near $(x = 0, t = 0)$. This is due to the fact that the function $\delta(x)$ is not exactly equal to the Dirac function. We correct the error as follows. It follows from (2.2), (2.7) and (2.8) that $u(x, t) = 1/2$ in a neighborhood of the point $(x, t) = (0, 0)$. We, therefore, replace the data $u(x, t)$ by $1/2$ when $|t|$ is small. In our computation, we set $u(x, t) = 1/2$ for $(x, t) \in [0, 0.0067] \times [0, 0.26]$. This data correction step is exactly the same as in section 6.1 of [24] and is illustrated by figure 2 in that publication. Then, we can extract the noiseless data $g_0$ easily. We next add the noise into the data via the formula

$$g_0 = g_0(1 + \delta \cdot \text{rand}), \quad (10.2)$$

where $\delta$ is the noise level and rand is the function that generates uniformly distributed random numbers in the range $[-1, 1]$. In all numerical tests with simulated data below, the noise level $\delta = 0.05$, i.e. $5\%$. Due to (2.12), the function $g_1 = g_0$. Due to the presence of noise, see (10.2), we cannot compute $g_1 = g_0$ by the finite difference method. Hence, the function $g_0$ is computed by the Tikhonov regularization method. The version of the Tikhonov regularization method for this problem is well-known. Hence, we do not describe this step here.

Having the data for the function $q$ in hand, we proceed as in algorithm 1.

In step 1 of algorithm 1, we choose $\lambda = 2$, $\alpha = 0.3$ and $\beta = 10^{-11}$. These parameters were chosen by a trial-error process that is similar to the one in [40]. Just as in [40], we choose a reference numerical test in which we know the true solution. In fact, test 1 of section 10.2 was our reference test. We have tried several values of $\lambda$, $\alpha$, and $\beta$ until the numerical solution to that reference test became acceptable. Then, we have used the same values of these parameters for all other tests, including all five (5) available cases of experimental data.

We next implement steps 2 and 3 of algorithm 1. We write differential operators in the functionals $J_{\lambda,\beta}^{(0)}$ and $J_{\lambda,\beta}^{(n)}$ in the finite differences with the step size in space $\Delta x = 0.0033$ and the step size in time $\Delta t = 0.02$ and minimize resulting functionals with respect to values of corresponding functions at grid points. Since the integrand in the definitions of the functional $J_{\lambda,\beta}^{(n)}$, $n = 0, 1, \ldots$ is the square of linear functions, then its minimizer is its critical point. In
finite differences, we can write a linear system whose solution is the desired critical point. We solve this system by the command ‘lsqlin’ of Matlab. The details in implementation by the finite difference method including the discretization, the derivation of the linear system for the critical point, and the use of ‘lsqlin’ are very similar to the scheme in [35]. Recall that in our theory, in the definition of the functional $J_{\lambda, \beta}$ acting on $q_n$, see (4.8), we replaced $q_{n-1}$ with its analog $q_{n-1,min}$ which belongs to the bounded set $B(R, g_0)$, and also replaced $q_{n-1,min}(x, 0)$ with $q_{n-1,min}^0(x, 0)$. These replacements are only for the theoretical purpose to avoid the case when $q_{n-1}$ blows up. However, in the numerical studies, these steps can be relaxed. This means that in step 3, we have minimized the finite difference analog of the functional

$$q_n \mapsto \int_\Omega \left( q_{n,x}(x, t) - \frac{q_{n,x}(x, t)}{2(q_{n-1}(x, 0))^2} + \frac{\partial_t q_{n-1}(x, t) \partial_t q_{n-1}(x, 0)}{2(q_{n-1}(x, 0))^2} \right)^2 \times e^{2\lambda g} \, dx \, dt + \beta \|q_n\|^2_{L^2(\Omega)},$$

subject to the boundary conditions in lines 2 and 3 of (3.11). Another numerical simplification is that rather than using the $H^4$-norm in the regularization term, we use the $H^2$-norm in (10.3). Although the theoretical analysis supporting the above simplifications is missing, we did not experience any difficulties in computing the numerical solutions to problem 3.1. All of our numerical results are satisfactory.

### 10.2. Numerical results for computationally simulated data

To test algorithm 1, we present four (4) numerical examples.

**Test 1 (the reference test).** We first test the case of one inclusion with a high inclusion/background contrast. The true dielectric constant function $c(x)$ has the following form

$$c_{true}(x) = \begin{cases} 
1 + 14 e^{\frac{(x-1)^2}{0.2}} & \text{if } |x - 1| < 0.2, \\
1 & \text{otherwise}.
\end{cases}$$

(10.4)

Thus, the inclusion/background contrast as 15:1. It is evident from figure 1 that we can successfully detect an object. The diameter of this object is 0.4 and the distance between this object and the source is 1. The true inclusion/background contrast is 15:1. The maximal value of the computed dielectric constant is 15.28. The relative error in this maximal value is 1.89% while the noise level in the data is 5%. Although the contrast is high, our method provides good numerical solution without requiring any knowledge of the true solution. Our method converges fast. Although the initial reconstruction $c_{init}$ computed in step 2 of algorithm 1 is not very good, see figure 1(a), one can see in figure 1(b), that the convergence occurs after only five (5) iterations. This fact verifies numerically the ‘contraction property’ of theorem 8.1 including the key estimate (8.20).

**Test 2.** The true function $c$ in this test has two ‘inclusions’,

$$c_{true}(x) = \begin{cases} 
1 + 5 e^{\frac{|x-0.6|^2}{0.2}} & \text{if } |x - 0.6| < 0.2, \\
1 + 8 e^{\frac{|x-1.4|^2}{0.2}} & \text{if } |x - 1.4| < 0.3, \\
1 & \text{otherwise}.
\end{cases}$$

(10.5)

Numerical results of this test are displayed in figure 2.
Test 2 is more complicated than test 1. However, we still obtain good numerical results. It is evident from figure 2(a) that we can precisely detect the two inclusions without any initial guess. The true maximal values of the dielectric constant of the inclusions in the left and the right are 6 and 9 respectively. The reconstructed maximal values in those inclusions are acceptable. They are 5.31 (relative error 11.5%) and 7.8 (relative error 13.3%). As in test 1, the initial reconstruction \( c_{\text{init}} \) computed in step 2 of algorithm 1 is far from \( c_{\text{true}}(x) \). Still, our iterative procedure converges after 7 iterations, see figure 2(b).
Test 3. We test the case when the function \( c_{\text{true}}(x) \) is discontinuous. It is a piecewise constant function given by

\[
c_{\text{true}}(x) = \begin{cases} 
10 & \text{if } |x - 1| < 0.15, \\
1 & \text{otherwise}.
\end{cases}
\]  
(10.6)

The numerical solution for this test is presented in figure 3.

Although the function \( c_{\text{true}} \) is not smooth and actually not even continuous, algorithm 1 works and provides a reliable numerical solution. The computed maximal value of the dielectric constant of the object is 9.25 (relative error 7.5%), which is acceptable. The location of the object is detected precisely, see figure 3(a). As in the previous two tests, algorithm 1 converges fast. After the fifth iteration, the reconstructed function \( c_n \) becomes unchanged. Again, this fact numerically confirms both theorem 8.1 and the robustness of our method.

Test 4. In this test, the function \( c_{\text{true}}(x) \) has the following form:

\[
c_{\text{true}}(x) = \begin{cases} 
3.5 + 0.3 \cdot \sin(\pi(x - 1.35)) & \text{if } |x - 0.9| < 0.5, \\
8 & \text{if } |x - 2| < 0.37, \\
1 & \text{otherwise}.
\end{cases}
\]  
(10.7)

This test is interesting since the graph of the function (10.7) consists of a curve and an inclusion. The numerical solution for this case is presented in figure 4.

One can observe from figure 4(a) that our method to compute the initial reconstruction in step 2 of algorithm 1 is not very effective. However, after only 6 iterations, good numerical results are obtained. The curve in the first inclusion locally coincides with the true one and the maximal value of the computed dielectric constant within inclusion is quite accurate: it is 7.83 (relative error 2.12%). Our method converges at the iteration number 6.

Remark 10.1. It follows from all above tests that algorithm 1 is robust in solving a highly nonlinear and severely ill-posed problem 3.1. It provides satisfactory numerical results with a
Figure 4. Test 4. The true and reconstructed functions $c(x)$, where $c_{\text{true}}$ is given in (10.7).
(a) The functions $c_{\text{init}}$ and $c_{\text{comp}}$ are obtained by steps 2 and 5 of algorithm 1 respectively. 
(b) The consecutive relative error $\|c_n - c_{n-1}\|_{L^\infty(\epsilon, M)}/\|c_n\|_{L^\infty(\epsilon, M)}$, $n = 1, \ldots, 10$. The data is with $\delta = 5\%$ noise.

fast convergence. For each test, the computational time to compute the numerical solution is about 29 seconds on a MacBook Pro 2019 with 2.6 GHz processor and 6 Intel i7 cores. This is almost a real time computation.

11. Numerical results for experimental data

We now test algorithm 1 on experimental data mentioned in introduction. These data were collected by the US Army Research Laboratory to detect and identify targets mimicking shallow anti-personnel land mines and IEDs. Five tested targets were: a bush, a wood stake, a metal box, a metal cylinder, and a plastic cylinder. The bush and the wood stake were placed in the air, while the other three objects were buried at a few centimeters depth in the ground. Since the locations of targets can be accurately detected by the Ground Position System, we are only interested in computing the values of their dielectric constants. We are doing so using algorithm 1.

Just as in our earlier works [15, 21, 22, 24, 28, 39], where these experimental data were used, we compute here the function $c_{\text{rel}}(x)$ defined as:

$$c_{\text{rel}}(x) = \begin{cases} 
\frac{c_{\text{target}}}{c_{\text{bckgr}}} & \text{if } \max \frac{c_{\text{target}}}{c_{\text{bckgr}}} (x) > 1 \text{ and } x \in D, \\
1 & \text{otherwise,}
\end{cases}$$

$$c_{\text{rel}}(x) = \begin{cases} 
\min \frac{c_{\text{target}}}{c_{\text{bckgr}}} & \text{if } \max \frac{c_{\text{target}}}{c_{\text{bckgr}}} (x) \leq 1 \text{ and } x \in D, \\
1 & \text{otherwise,}
\end{cases}$$

where $D$ is a sub-interval of the interval $[\epsilon, M]$, which is occupied by the target. Next, we define the computed value of $c_{\text{target}}$ as [24]:

$$c_{\text{comp}} = c_{\text{bckgr}} \times \begin{cases} 
\max c_{\text{rel}} & \text{if } \max c_{\text{rel}} > 1, \\
\min c_{\text{rel}} & \text{if } \max c_{\text{rel}} < 1.
\end{cases}$$
As in the above cited publications, we have to preprocess the raw data of [34] before importing them into our solver. The data preprocessing procedure is exactly the same as the one in [24, section 7.1]. First, we observe that the $L_{\infty}$-norm of the experimental data far exceeds the $L_{\infty}$-norm of the computationally simulated data. Therefore, we need to scale the experimental data by dividing it by a calibration factor $\mu > 0$, i.e. we replace the raw experimental data $f_{\text{raw}}(t)$ with $f_{\text{scale}}(t) = f_{\text{raw}}(t)/\mu$. Then we work only with $f_{\text{scale}}(t)$. To find the calibration factor, we use a trial-and-error process. First, we select a reference target. We then try many values of $\mu$ such that the reconstruction of the reference target is satisfactory, i.e. the computed dielectric constant is in its published range, see below in this section about the published range. Then this calibration factor is used in all other tests. When the object is in the air, our reference target is bush. In this case, the calibration factor $\mu_{\text{air}} = 534,592$. When the object is buried under the ground, our reference target is the metal box and the calibration factor was $\mu_{\text{ground}} = 265,223$.

Next, we preprocess the data $f_{\text{scale}}(t)$ as follows. First, we first use a lower envelop (built in Matlab) to bound the data. We then truncate the data that is not in a small interval centered at the global minimizer of the data, see [24, section 7.1] for the choice of this small interval. But in the case of the plastic cylinder we use the upper envelop. The choice of the upper or lower envelopes is as follows. We look at the function $f_{\text{scale}}(t)$ and find the three extrema with largest absolute values. If the middle extremal value among these three is a minimum, then we bound the data by a lower envelop. Otherwise, we use an upper envelop. See [24, section 7.1] for more details and the reason of this choice. In particular, figures 4(a)–(d), 5(a), (b), (d), (e), (g) and (h) of [24] provide illustrations. Likewise, our figure 5 displays computed functions $c_{\text{target}}(x,y)$ for our five targets. The computed dielectric constants $c_{\text{comp}}$ defined in (11.3) by algorithm 1 are listed in table 1.

The true values of dielectric constants of our targets were not measured in the experiments. Therefore, we compare our computed values with the published ones. The published values of the dielectric constants of our targets are listed in the last column of table 1. They can
Table 1. Computed dielectric constants of five targets.

| Target        | $c_{\text{backg}}$ | Computed $c_{\text{rel}}$ | $c_{\text{backg}}$ | Computed $c_{\text{target}}$ | True $c_{\text{target}}$ |
|---------------|---------------------|----------------------------|---------------------|-------------------------------|---------------------------|
| Bush          | 1                   | 7.62                       | 1                   | 7.62                          | [3, 20]                   |
| Wood stake    | 1                   | 2.01                       | 1                   | 2.01                          | [2, 6]                    |
| Metal box     | 4                   | 4.00                       | [3, 5]              | [12.00, 20.00]                | [10, 30]                  |
| Metal cylinder| 4                   | 4.01                       | [3, 5]              | [12.3, 20.5]                  | [10, 30]                  |
| Plastic cylinder | 4              | 0.59                       | [3, 5]              | [1.6, 2.95]                   | [1.1, 3.2]                |

be found on the website of Honeywell (table of dielectric constants, https://goo.gl/kAxtzB). Also, see [11] for the experimentally measured range of the dielectric constants of vegetation samples, which we assume have the same range as the dielectric constant of bush. In the table of dielectric constants of Honeywell as well as in [11], any dielectric constant is not a number. Rather, each dielectric constant of these references is given within a certain interval. As to the metallic targets, it was established in [28] that they have the so-called ‘apparent’ dielectric constant whose values are in the interval [10, 30].

**Conclusion.** It is clear from table 1 that our computed dielectric constants are consistent with the intervals of published ones. Therefore, our results for experimental data are satisfactory.

**Remark 11.1 (Speed of computations).** Our experimental data are sparse. The size of the data in time is $N_t = 80$, which is a lot smaller than that in the dense simulated data ($N_t = 300$). Therefore, the speed of computations is much faster than for the case of simulated data of section 10. All results of this section were computed in real time on the same computer (MacBook Pro 2019 with 2.6 GHz processor and 6 Intel i7 cores).

12. **Concluding remarks**

We have developed a new globally convergent numerical method for a 1D CIP with backscattering data for the wave-like PDE (2.3). This is the second generation of the above cited convexification method of our research group. The main novelty here is that, rather than minimizing a globally strictly convex weighted cost functional arising in the convexification, we solve on each iterative step a linear BVP. This is done using the so-called CQRM. Just like in the convexification, the key element of the convergence analysis of this paper is the presence of the CWF in each quadratic functional, which we minimize. The convergence estimate is similar to the well known estimate of the classical contraction mapping principle. The latter explains the title of this paper. We have proven a global convergence theorem of our method. Our numerical results for computationally simulated data demonstrate high reconstruction accuracies in the presence of 5% random noise in the data.

Furthermore, our numerical results for experimentally collected data have satisfactory accuracy via providing values of computed dielectric constants of explosive-like targets within their published ranges.

A practically important observation here is that our computations for experimental data were performed in real time. This observation did not take place for various versions of the first generation of the convexification method of our previous publications [21, 22, 24, 39], which were working with the same experimental data. The latter indicates an important advantage of the second generation of the convexification method of this paper.
Acknowledgments

The effort of TTL, MVK and LHN was supported by US Army Research Laboratory and US Army Research Office Grant W911NF-19-1-0044.

Data availability statement

The data that support the findings of this study are available upon reasonable request from the authors.

ORCID iDs

Michael V Klibanov  https://orcid.org/0000-0001-5179-6879
Loc H Nguyen  https://orcid.org/0000-0002-0172-8816

References

[1] Bakushinskii A B, Klibanov M V and Koshev N A 2017 Carleman weight functions for a globally convergent numerical method for ill-posed Cauchy problems for some quasilinear PDEs Nonlinear Anal.: Real World Appl. 34 201–24
[2] Baudouin L, de Buhan M and Ervedoza S 2013 Global Carleman estimates for waves and applications Commun. PDE 38 823–59
[3] Baudouin L, de Buhan M and Ervedoza S 2017 Convergent algorithm based on Carleman estimates for the recovery of a potential in the wave equation SIAM J. Numer. Anal. 55 1578–613
[4] Baudouin L, de Buhan M, Ervedoza S and Osses A 2021 Carleman-based reconstruction algorithm for the waves SIAM J. Numer. Anal. 59 998–1039
[5] Beilina L and Klibanov M V 2012 Approximate Global Convergence and Adaptivity for Coefficient Inverse Problems (New York: Springer)
[6] Bellassoued M and Yamamoto M M 2017 Carleman Estimates and Applications to Inverse Problems for Hyperbolic Systems (Berlin: Springer)
[7] Bourgeois L 2006 Convergence rates for the quasi-reversibility method to solve the Cauchy problem for Laplace’s equation Inverse Problems 22 413–30
[8] Bourgeois L and Dardé J 2010 A duality-based method of quasi-reversibility to solve the Cauchy problem in the presence of noisy data Inverse Problems 26 095016
[9] Bukhgeim A L and Klibanov M V 1981 Uniqueness in the large of a class of multidimensional inverse problems Sov. Math. - Dokl. 17 244–7
[10] Chavent G 2009 Nonlinear Least Squares for Inverse Problems: Theoretical Foundations and Step-by-Step Guide for Applications Scientific Computation (New York: Springer)
[11] Chua H T, Lee K Y and Lau T W 1995 Dielectric constants of rubber and oil palm leaf samples at X-band IEEE Trans. Geosci. Remote Sens. 33 221–3
[12] de Hoop M V, Kepley P and Oksanen L 2018 Recovery of a smooth metric via wave field and coordinate transformation reconstruction SIAM J. Appl. Math. 78 1931–53
[13] Goncharsky A V and Romanov S Y 2017 Iterative methods for solving coefficient inverse problems of wave tomography in models with attenuation Inverse Problems 33 025003
[14] Goncharsky A V and Romanov S Y 2019 A method of solving the coefficient inverse problems of wave tomography Comput. Math. Appl. 77 967–80
[15] Karchevsky A L, Klibanov M V, Nguyen L, Pantong N and Sullivan A 2013 The Krein method and the globally convergent method for experimental data Appl. Numer. Math. 74 111–27
[16] Klibanov M V 1992 Inverse problems and Carleman estimates Inverse Problems 8 575–96
[17] Klibanov M V 2013 Carleman estimates for global uniqueness, stability and numerical methods for coefficient inverse problems J. Inverse Ill-Posed Problems 21 477–560
[18] Klibanov M V 2015 Carleman estimates for the regularization of ill-posed Cauchy problems Appl. Numer. Math. 94 46–74
[19] Klibanov M V and Joussoupova O V 1995 Uniform strict convexity of a cost functional for three-dimensional inverse scattering problem *SIAM J. Math. Anal.* **26** 147–79

[20] Klibanov M V 1997 Global convexity in a three-dimensional inverse acoustic problem *SIAM J. Math. Anal.* **28** 1371–88

[21] Klibanov M V, Kolesov A E, Nguyen L and Sullivan A 2017 Globally strictly convex cost functional for a 1D inverse medium scattering problem with experimental data *SIAM J. Appl. Math.* **77** 1733–55

[22] Klibanov M V, Kolesov A E, Nguyen L and Sullivan A 2018 A new version of the convexification method for a 1D coefficient inverse problem with experimental data *Inverse Problems* **34** 115014

[23] Klibanov M V, Li J and Zhang W 2020 Convexification for an inverse parabolic problem *Inverse Problems* **36** 085008

[24] Klibanov M V, Le T T, Nguyen L H, Sullivan A and Nguyen L 2021 Convexification-based globally convergent numerical method for a 1D coefficient inverse problem with experimental data *Inverse Problems Imaging* https://doi.org/10.3934/ipi.2021068

[25] Klibanov M V and Li J 2021 *Inverse Problems and Carleman Estimates: Global Uniqueness, Global Convergence and Experimental Data* (Berlin: de Gruyter & Co)

[26] Klibanov M V, Khoa V A, Smirnov A V, Nguyen L H, Bidney G W, Nguyen L, Sullivan A and Astratov V N 2021 Convexification inversion method for nonlinear SAR imaging with experimentally collected data *J. Appl. Ind. Math.* **15** 413–36

[27] Korpela J, Lassas M and Oksanen L 2016 Regularization strategy for an inverse problem for a 1 + 1 dimensional wave equation *Inverse Problems* **32** 065001

[28] Kuzhuget A V, Beilina L, Klibanov M V, Sullivan A, Nguyen L and Fiddy M A 2012 Blind backscattering experimental data collected in the field and an approximately globally convergent inverse algorithm *Inverse Problems* **28** 095007

[29] Lattes R and Lions J L 1969 *The Method of Quasireversibility: Applications to Partial Differential Equations* (New York: Elsevier)

[30] Lavrent’ev M M, Romanov V G and Shishatskii S P 1986 *Ill-Posed Problems of Mathematical Physics and Analysis* (Providence, RI: American Mathematical Society)

[31] Le T T and Nguyen L H 2020 A convergent numerical method to recover the initial condition of nonlinear parabolic equations from lateral Cauchy data *J. Inverse Ill-Posed Problems* https://doi.org/10.1515/jiip-2020-0028

[32] Levitan B M 1987 *Inverse Sturm–Liouville Problems* (Utrecht: VSP)

[33] Minoux M 1986 *Mathematical Programming: Theory and Algorithms* (New York: Wiley)

[34] Nguyen N, Wong D, Ressler M, Koenig F, Stanton B, Smith G, Sichina J and Kappra K 2007 Obstacle avoidance and concealed target detection using the Army Research Lab ultra-wideband synchronous impulse reconstruction (UWB SIRE) forward imaging radar *Proc. SPIE* **6553** 65530H

[35] Nguyen L H 2020 A new algorithm to determine the creation or depletion term of parabolic equations from boundary measurements *Comput. Math. Appl.* **80** 2135–49

[36] Romanov V G 1986 *Inverse Problems of Mathematical Physics* (Utrecht: VNU Press)

[37] Romanov V G 2002 *Investigation Methods for Inverse Problems* (Utrecht: VSP)

[38] Scales J A, Smith M L and Fischer T L 1992 Global optimization methods for multimodal inverse problems *J. Comput. Phys.* **103** 258–68

[39] Smirnov A V, Klibanov M V, Sullivan A J and Nguyen L H 2020 Convexification for an inverse problem for a 1D wave equation with experimental data *Inverse Problems* **36** 095008

[40] Smirnov A, Klibanov M and Nguyen L 2020 Convexification for a 1D hyperbolic coefficient inverse problem with single measurement data *Inverse Problems Imaging* **14** 913–38

[41] Tikhonov A N, Goncharsky A V, Stepanov V V and Yagola A G 1995 *Numerical Methods for the Solution of Ill-Posed Problems* (Dordrecht: Kluever)

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