Convergent numerical methods for parabolic equations with reversed time via a new Carleman estimate

Michael V Klibanov\textsuperscript{1,3} and Anatoly G Yagola\textsuperscript{2}

\textsuperscript{1} Department of Mathematics and Statistics, University of North Carolina Charlotte, Charlotte, NC 28223, United States of America
\textsuperscript{2} Faculty of Physics, Department of Mathematics, Moscow State University, Moscow, 119991, Russia

E-mail: mklibanv@uncc.edu and yagola@physics.msu.ru

Received 27 February 2019, revised 13 May 2019
Accepted for publication 6 June 2019
Published 4 October 2019

Abstract
The key tool of this paper is a new Carleman estimate for an arbitrary parabolic operator of the second order for the case of reversed time data. This estimate works on an arbitrary time interval. On the other hand, the previously known Carleman estimate for the reversed time case works only on a sufficiently small time interval. First, a stability estimate is proven. Next, the quasi-reversibility numerical method is proposed for an arbitrary time interval for the linear case. This is unlike a sufficiently small time interval in the previous work. The convergence rate for the quasi-reversibility method is established. Finally, the quasilinear parabolic equation with reversed time is considered. A weighted globally strictly convex Tikhonov-like functional is constructed. The weight is the Carleman weight function which is involved in that Carleman estimate. The global convergence of the gradient projection method to the exact solution is proved for this functional.

Keywords: linear and quasilinear parabolic equations, reversed time, Carleman estimate, stability estimate, convergent quasi-reversibility method for the linear case, globally convergent numerical method for the quasilinear case

1. Introduction

In this paper, we construct convergent numerical methods for linear and quasilinear parabolic equations with reversed time. The key tool of the convergence analysis is a new Carleman estimate. While the previously known Carleman estimate for these problems works only on a sufficiently small time interval, the one of this paper works on any finite time interval.

\textsuperscript{3} Author to whom any correspondence should be addressed.
One of possible applications of problems considered here is in the case when a solid is heated and the initial temperature is unknown. However, one can measure the temperature of this solid at a final time. It is required to restore the temperature distribution inside of this solid at all preceding times. Another recently found application of this problem is in the financial mathematics. More precisely, in the problem of forecasting of prices of stock options using the Black–Scholes equation and real market data [19].

All functions below are real valued ones. Below \( x = (x_1, x_2, ..., x_n) \) denotes points in \( \mathbb{R}^n \) and \( \nabla f = (f_1, f_2, ..., f_n) \) for any appropriate function \( f(x) \). Let \( \Omega \subset \mathbb{R}^n \) be a bounded domain with a piecewise smooth boundary \( \partial \Omega \). Let \( T > 0 \) and \( \tau \in (0, T) \) be two numbers. Denote

\[
Q_T = \Omega \times (0, T), S_T = \partial \Omega \times (0, T), Q_{T_\tau} = \Omega \times (\tau, T).
\]

For \( i, j = 1, ..., n \), let functions \( a_{ij}(x, t) \) be such that

\[
a_{ij}(x, t) = a_{ij}(x, t) \in \mathcal{C}^1(Q_T),
\]

\[
\mu_1 |\xi|^2 \leq \sum_{i,j=1}^n a_{ij}(x, t) \xi_i \xi_j \leq \mu_2 |\xi|^2, \forall (x, t) \in Q_T, \forall \xi \in \mathbb{R}^n,
\]

\[
\mu_1, \mu_2 = \text{const.} > 0, \mu_1 \leq \mu_2.
\]

Introduce a uniformly elliptic operator \( L \) of the second order in the domain \( Q_T \),

\[
Lu = \sum_{i,j=1}^n a_{ij}(x, t) u_{x_i x_j}, \quad (x, t) \in Q_T.
\]

Let the function \( F(y) \in \mathcal{C}^1(\mathbb{R}^{2n+2}) \). For this function, we assume that there exists a constant \( \overline{C} = \overline{C}(F) > 0 \) depending only on \( F \) such that

\[
|F(y_1) - F(y_2)| \leq \overline{C}|y_1 - y_2|, \forall y_1, y_2 \in \mathbb{R}^{2n+2}.
\]

Consider the following quasilinear parabolic equation:

\[
\begin{aligned}
&u_t = Lu + F(\nabla u, u, x, t), \quad (x, t) \in Q_T. \\
&u|_{\partial Q_T} = 0.
\end{aligned}
\]

We impose the zero Dirichlet boundary condition on the function \( u \),

\[
(1.8)
\]

Since we work with the time reversed case, we assume that the function \( u(x, t) \) is known at the final time \( T \),

\[
(1.9)
\]

Thus, we have obtained the following problem:

**Problem with time reversed data.** Suppose that conditions (1.1)–(1.7) hold. Find a function \( u \in \mathcal{C}^{2,1}(\overline{Q_T}) \) satisfying conditions (1.7)–(1.9).

The problem of our interest is well known to be unstable, i.e. this is an ill-posed problem. Hölder stability estimates for this problem are known, see, e.g. [10, 16, 26]. To the best knowledge of the author, the strongest Hölder stability result, which is valid for an arbitrary large time interval \( t \in (0, T) \), is obtained by Isakov, see theorem 3.1.3 in [10]. In section 2 of chapter 4 of [26] and later in [16] a Carleman estimate was used to obtain the Hölder stability estimate. However, that Hölder stability estimate is valid only on a sufficiently small time interval \( t \in (T - \varepsilon, T) \) for a sufficiently small \( \varepsilon > 0 \). The smallness of this interval is due to the Carleman weight function (CWF), which has been used in the Carleman estimates for that
problem so far [16, 26]. This function is $(k + T - t)^{-2\lambda}$, where numbers $k, T > 0$ are sufficiently small and the parameter $\lambda > 0$ is sufficiently large. The same CWF was used in [8] for the proof of the uniqueness theorem.

Using this estimate, the first author has constructed in [16] the quasi-reversibility method (QRM) for the above problem in the linear case and has proven its convergence, again on a sufficiently small time interval. Below as well as in [16, 19] the QRM is realized via the minimization of a certain regularization functional. On the other hand, the QRM is quite often realized via a proper perturbation of the underlying PDE operator [10, 27]. A surprising idea of the recent publication of Kaltenbacher and Rundell [11] it to use the non local operator of the fractional $t$-derivative as the perturbing operator for the QRM. In the work of Tuan, Khoa and Au [30] another version of the QRM is constructed for the quasilinear case. Its convergence was proven for an arbitrary $T$. However, the perturbation operator of [30] is a very complicated one. Another version of the QRM was proposed in [10]. This version works only in the case when the set of eigenfunctions of the underlying elliptic operator forms an orthonormal basis in $L^2(\Omega)$.

We now briefly compare results of this paper with the ones of the paper of Hao et al [9]. In [9] the problem with the time reversed data is considered for the quasilinear equation $u_t(t) + A(t)u(t) = f(t, u(t))$, where $A(t)$ is a positive self-adjoint unbounded operator in a Hilbert space $H$. This is close to the problem we consider here, although we do not assume that our operator $L$ in (1.5) is self-adjoint. The function $f(t, u(t))$ in [9] satisfies the local Lipschitz condition with respect to $u$, which is more general than our global Lipschitz condition (1.6). The applied importance of the local Lipschitz condition is demonstrated in [9] using some specific examples of parabolic PDEs arising in Physics. A Hölder stability estimate is proven in [9]. This estimate of [9] is stronger than ours, since our stability estimate of theorem 3.1 is ‘between’ Hölder and logarithmic types: see item 2 below as well as remark 3.1 in section 3. The Tikhonov functional is constructed in [9]. The difference between the minimizer of that functional and the exact solution of the original problem is estimated in [9] using the above mentioned Hölder stability estimate. In the quasilinear case, our weighted Tikhonov-like functional is different from the one of [9].

The QRM was originally introduced by Lattes and Lions in 1969 [27]. Their idea became quite popular since then with many publications treating a variety of ill-posed problems for PDEs. In this regard we refer to, e.g. [5–7, 10, 12, 19, 25, 30]. In particular, the first author has shown in the survey paper [16] that as soon as a proper Carleman estimate for an ill-posed problem for a linear PDE is available, then the QRM can be constructed for this problem and its convergence rate can be established. However, in the case of the time reversed data for the linear parabolic PDE the construction of the QRM in [16] is valid only on a sufficiently small time interval $(0, T)$.

Ill-posed problems for quasilinear PDEs are nonlinear ones. It is well known that a conventional Tikhonov functional for a nonlinear ill-posed problem is non convex. Hence, one cannot guarantee the absence of local minima and ravines of this functional, see, e.g. [28] for a numerical example. Therefore, it is desirable to construct such numerical methods for ill-posed problems for quasilinear PDEs, which would not suffer from the phenomenon of local minima and ravines. This question was addressed in [17] for a class of ill-posed problems for quasilinear PDEs. It was established in [17] that, given an ill-posed problem for a quasilinear PDE, if a proper Carleman estimate is available for the linear principal part of that PDE operator, then a weighted globally strictly convex Tikhonov-like functional can be constructed for that problem, i.e. this problem can be ‘convexified’. 
The key element of this functional is the presence of the CWF in it. This is the function which is involved as the weight in the Carleman estimate for the linear principal part of that PDE operator. In the follow up paper [1] existence and uniqueness of the minimizer of that functional were established and global convergence to the exact solution of the gradient projection method of the minimization of this functional was proven. In both works [1, 17] examples of ill-posed problems for quasilinear parabolic, elliptic and hyperbolic PDEs were provided. As to the quasilinear parabolic equations, the technique of [17] was applied in [1, 17, 20] only for the case of the lateral Cauchy data. In particular, see numerical results in [1, 20]. However, the case of the time reversed data for quasilinear parabolic equations was not yet handled by the technique of [17]. Thus, using our new Carleman estimate, we construct the desired method in section 5.

The idea of convexifying coefficient inverse problems was originally proposed in 1995–1997 in [13, 14], although without numerical studies, also see [15]. Recently the interest to the convexification approach was renewed. Initially this was done only analytically in [4, 18]. Next, a number of papers was published, which combined the theory with numerical studies, see, e.g. [1, 20] for the case of quasilinear PDEs and [2, 21–24] for coefficient inverse problems. In particular, experimentally collected data were treated by the convexification in [21, 22].

We call a numerical method for an ill-posed problem globally convergent if there is a theorem claiming that it converges to the exact solution of this problem without an a priori knowledge of a sufficiently small neighborhood of this solution. Thus, the convexification is a globally convergent numerical method (see theorem 5.4 in section 5).

New elements of this paper are:

1. A new Carleman estimate for a general parabolic operator of the second order with time reversed data is proven. This estimate works on an arbitrary time interval \( t \in (0, T) \), unlike a sufficiently small interval of previous publications [8, 16, 26]. Results listed in items 2–4 below are based on this estimate.

2. A stability estimate is proven for the above mentioned problem with reversed time. This estimate is somewhat ‘between’ Hölder and logarithmic stability estimates. In other words, it is weaker than the Hölder stability estimate of [10, 16, 26] but it is stronger than the logarithmic stability estimate. Still, the main advantage of our stability estimate over ones in [16, 26] is that it works without a smallness assumption imposed on the time interval.

3. In the linear case, the QRM is constructed, existence and uniqueness of the minimizer as well as convergence of minimizers to the exact solution are proven. Unlike previous works [16, 19], a smallness assumption is not imposed on the time interval.

4. The problem with time reversed data for a quasilinear parabolic PDE is convexified. For the constructed weighed Tikhonov-like functional with a CWF in it, both the existence and uniqueness of its minimizer are proved. In addition, it is established that the gradient projection method, being applied to this functional, converges globally to the exact solution. Thus, we do not face the above discussed problem of local minima and ravines. Also, we avoid the use of a complicated perturbation operator of [30].

In section 2 we prove the new Carleman estimate. In section 3 we present stability estimate. In section 4 we describe the QRM for the linear case, prove existence and uniqueness of the minimizer of the corresponding functional as well as convergence of minimizers to the exact solution when the level of the noise in the data tends to zero. In section 5 we construct the above mentioned weighted globally strictly convex Tikhonov-like functional and formulate corresponding theorems. We prove these theorems in section 6.
2. Carleman estimate

Theorem 2.1 (Carleman estimate). Assume that conditions (1.1)–(1.5) are in place. Then there exist numbers

\[
\nu_0 = \nu_0 \left( \mu_1, \mu_2, \max_{ij} \| a_{ij} \|_{C^1(\overline{Q}_T)}, Q_T \right) > 1,
\]

\[
C = C \left( \mu_1, \mu_2, \max_{ij} \| a_{ij} \|_{C^1(\overline{Q}_T)}, Q_T \right) > 0
\]

(2.1)
depending only on listed parameters, \( \nu_0 \) sufficiently large, such that for all functions \( u \in H^2(Q_T) \) satisfying the zero Dirichlet boundary condition (1.8) the following Carleman estimate is valid

\[
\int_{Q_T} (u_t - Lu)^2 \exp \left( 2 (t + 1)\nu \right) \, dx \, dt
\]

\[
\geq \frac{\mu_1}{6} \sqrt{\nu} \int_{Q_T} (\nabla u)^2 \exp \left( 2 (t + 1)\nu \right) \, dx \, dt + \frac{\nu^2}{12} \int_{Q_T} u^2 \exp \left( 2 (t + 1)\nu \right) \, dx \, dt
\]

\[
- \exp \left( 3 (T + 1)\nu \right) \| u(x, T) \|_{L^2(\Omega)}^2 - \mu_2 \nu^2 \| \nabla u(x, 0) \|_{L^2(\Omega)}^2, \quad \forall \nu \geq \nu_0.
\]

(2.2)

Proof. Everywhere below in this paper \( C > 0 \) denotes different constants depending only on parameters listed in (2.1).

In this proof,

\[
u \in C^2(\overline{Q}_T), u \mid_{\partial T} = 0.
\]

(2.3)

The case \( u \in H^2(Q_T), u \mid_{\partial T} = 0 \) can be obtained from (2.3) via density arguments.

We prove this theorem in six steps. Introduce a new function \( v(x, t) \),

\[
v(x, t) = u(x, t) \exp \left( (t + 1)\nu \right).
\]

Hence,

\[
u_t = \nu_t - (t + 1)\nu - (t + 1)\nu \exp \left( (t + 1)\nu \right),
\]

\[
u_{xx} = \nu_{xx} \exp \left( (t + 1)\nu \right).
\]

Hence,

\[(u_t - Lu)^2 \exp \left( 2 (t + 1)\nu \right) = \left[ \nu_t - \left( L
\nu + (t + 1)\nu - (t + 1)\nu \right) \right]^2
\]
\[ \geq v_t^2 - 2v_t \left( LV + \nu (t + 1)^{\nu-1} v \right) \]  \hspace{1cm} (2.4) \\
\[ = v_t^2 - 2v_t LV - 2\nu (t + 1)^{\nu-1} v_t v. \]

**Step 1.** First, we estimate from the below \(-2v_t LV\),

\[ -2v_t LV = -\sum_{i=1}^{n} \left( a_{i,j}v_{x_i}v_t + a_{i,j}v_{y_j}v_t \right). \]  \hspace{1cm} (2.5) 

Next,

\[ - (a_{i,j}v_{x_i}v_t + a_{i,j}v_{x_j}v_t) = - (a_{i,j}v_{x_i}v_t)_{x_i} + a_{i,j}v_{x_i}v_{v_t} + (a_{i,j})_{x_i} v_{x_i} v_t, \]

\[ - (a_{i,j}v_{x_j}v_t)_{x_i} + a_{i,j}v_{x_i}v_{v_t} + (a_{i,j})_{x_i} v_{x_i} v_t, \]

\[ = (a_{i,j}v_{x_i}v_t)_{x_i} - (a_{i,j})_{x_i} v_{x_i} v_t + (a_{i,j})_{x_i} v_{x_i} v_t \]

\[ - \left[ (a_{i,j}v_{x_i}v_t)_{x_i} + (a_{i,j}v_{x_i}v_t)_{x_i} \right]. \]

Applying the Cauchy–Schwarz inequality ‘with \(\varepsilon\)’ to (2.6),

\[ 2ab \geq -\varepsilon a^2 - \frac{1}{\varepsilon} b^2, \forall a, b \in \mathbb{R}, \forall \varepsilon > 0, \]  \hspace{1cm} (2.7) 

and using (2.5), we obtain

\[ -2v_t LV \geq -C (\nabla v)^2 - \frac{1}{2} v_t^2 + \sum_{i=1}^{n} \left( a_{i,j}v_{x_i}v_{v_t} \right)_{v_t}, \]

\[ -\sum_{i=1}^{n} \left[ (a_{i,j}v_{x_i}v_t)_{x_i} + (a_{i,j}v_{x_i}v_t)_{x_i} \right]. \]

Hence,

\[ v_t^2 - 2v_t LV \geq \frac{1}{2} v_t^2 - C (\nabla v)^2 \]

\[ + \left( \sum_{i=1}^{n} a_{i,j}v_{x_i}v_t \right)_{v_t} - \sum_{i=1}^{n} \left[ (a_{i,j}v_{x_i}v_t)_{x_i} + (a_{i,j}v_{x_i}v_t)_{x_i} \right] \]  \hspace{1cm} (2.8) 

\[ \geq -C (\nabla v)^2 + \left( \sum_{i=1}^{n} a_{i,j}v_{x_i}v_t \right)_{v_t} - \sum_{i=1}^{n} \left[ (a_{i,j}v_{x_i}v_t)_{x_i} + (a_{i,j}v_{x_i}v_t)_{x_i} \right]. \]

**Step 2.** Estimate the term \(-2\nu (t + 1)^{\nu-1} v_t v\) in the third line of (2.4). We have:

\[ -2\nu (t + 1)^{\nu-1} v_t v = \left( -\nu (t + 1)^{\nu-1} v_t^2 \right)_{v_t} + \nu (\nu - 1) (t + 1)^{\nu-2} v_t^2. \]  \hspace{1cm} (2.9)
Choose \( \nu_0 > 1 \) so large that
\[
\nu (\nu - 1) > \frac{\nu^2}{2}, \quad \forall \nu \geq \nu_0. \tag{2.10}
\]

**Step 3.** Sum up (2.8) and (2.9). Then, taking into account (2.4) and (2.10) and ignoring the term \( \nu_0^2/2 \geq 0 \), we obtain for \( \nu \geq \nu_0 \)
\[
(u_t - Lu)^2 \exp (2(t + 1)^\nu) \geq -C (\nabla v)^2 + \frac{\nu^2}{2} (t + 1)^{\nu-2} v^2
\]
\[
+ \left( \sum_{i,j=1}^{n} a_{ij} v_i v_j - \nu (t + 1)^{\nu-1} v^2 \right) + \sum_{i,j=1}^{n} \left[ (-a_{ij} v_i v_j)_x + (-a_{ij} v_i v_j)_y \right].
\]
Replacing here \( v \) with \( u = v \exp (-t + 1)^\nu \), we obtain
\[
(u_t - Lu)^2 \exp (2\lambda (t + 1)^\nu) \geq -C (\nabla u)^2 + \frac{\nu^2}{2} (t + 1)^{\nu-2} u^2 \exp (2(t + 1)^\nu)
\]
\[
+ \left( \sum_{i,j=1}^{n} a_{ij} u_i u_j - \nu (t + 1)^{\nu-1} u^2 \right) \exp (2(t + 1)^\nu) \right)_v
\]
\[
+ \sum_{i,j=1}^{n} \left[ -a_{ij} u_i \left( u_t + \nu (t + 1)^{\nu-1} u \right) \exp (2(t + 1)^\nu) \right]_{x_i}
\]
\[
+ \sum_{i,j=1}^{n} \left[ -a_{ij} u_i \left( u_t + \nu (t + 1)^{\nu-1} u \right) \exp (2(t + 1)^\nu) \right]_{x_j}.
\]
What is not good in estimate (2.11) is that the first line in its right hand side contains both positive and negative terms. However, only positive terms must be in such cases in any Carleman estimate. Hence, we proceed with further steps.

**Step 4.** Estimate from the below the expression \((u_t - Lu) u \exp (2(t + 1)^\nu)\). We have
\[
(u_t - Lu) u \exp (2(t + 1)^\nu) = \left( \frac{u^2}{2} \exp (2(t + 1)^\nu) \right) - \nu (t + 1)^{\nu-1} u^2 \exp (2(t + 1)^\nu)
\]
\[
- \frac{1}{2} \sum_{i,j=1}^{n} \left( a_{ij} u_i u_j \exp (2(t + 1)^\nu) \right)_v - \frac{1}{2} \sum_{i,j=1}^{n} \left( a_{ij} u_i u_j \exp (2(t + 1)^\nu) \right)_{x_i}
\]
\[
+ \sum_{i,j=1}^{n} a_{ij} u_i u_j \exp (2(t + 1)^\nu) \right)_{x_j} \tag{2.12}
\]
\[
+ \frac{1}{2} \sum_{i,j=1}^{n} \left( a_{ij} u_i u_j \exp (2(t + 1)^\nu) + \frac{1}{2} \sum_{i,j=1}^{n} \left( a_{ij} u_i u_j \exp (2(t + 1)^\nu) \right).
By (1.3)
\[\sum_{i,j=1}^{n} a_{ij}u_i u_j \exp (2(t+1)^\nu) \geq \mu_1 (\nabla u)^2 \exp (2(t+1)^\nu).\]  (2.13)

Next, using again the Cauchy–Schwarz inequality (2.7) \(\text{with } \epsilon'\), we obtain
\[
\begin{align*}
&+ \frac{1}{2} \sum_{i,j=1}^{n} (a_{ij})_{x_i u} u \exp (2(t+1)^\nu) + \frac{1}{2} \sum_{i,j=1}^{n} (a_{ij})_{x} u \exp (2(t+1)^\nu) \tag{2.14}
\end{align*}
\]
\[\geq -\frac{\mu_1}{2} (\nabla u)^2 \exp (2(t+1)^\nu) - Cu^2 \exp (2(t+1)^\nu).\]

Choose \(\nu_0 > 1\) so large that, in addition to (2.10),
\[
\nu_0 > C. \tag{2.15}
\]

Then we obtain from (2.12)–(2.15) for \(\nu \geq \nu_0\)
\[
(u_t - Lu) u \exp (2(t+1)^\nu) \geq \left[\frac{\mu_1}{4} (\nabla u)^2 - 2\nu (t+1)^{\nu-1} u^2\right] \exp (2(t+1)^\nu)
\]
\[+ \left(\frac{u^2}{2} \exp (2(t+1)^\nu)\right). \tag{2.16}
\]
\[- \frac{1}{2} \sum_{i,j=1}^{n} (a_{ij}u_i u \exp (2(t+1)^\nu))_{x_i} - \frac{1}{2} \sum_{i,j=1}^{n} (a_{ij}u_j u \exp (2(t+1)^\nu))_{x_j}.
\]

Choose \(\nu_0 > 1\) so large that, in addition to (2.10) and (2.15),
\[
\sqrt{\nu} > \max \left(\frac{4C}{\mu_1}, 8(T+1)\right), \forall \nu \geq \nu_0. \tag{2.17}
\]

**Step 5.** Multiply (2.16) by \(\sqrt{\nu}\) and sum up with (2.11). Using (2.17), we obtain for \(\nu \geq \nu_0\)
\[
(u_t - Lu)^2 \exp (2(t+1)^\nu) + \sqrt{\nu} (u_t - Lu) u \exp (2(t+1)^\nu)
\]
\[\geq \left[\frac{\mu_1}{4} \sqrt{\nu} (\nabla u)^2 + \frac{\nu^2}{4} (t+1)^{\nu-2} u^2\right] \exp (2(t+1)^\nu)
\]
\[+ \left[\left(\sum_{i,j=1}^{n} a_{ij}u_i u_j - \nu (t+1)^{\nu-1} u^2 + \sqrt{\nu} \frac{u^2}{2}\right) \exp (2(t+1)^\nu)\right]_{t}
\]
\[+ \sum_{i,j=1}^{n} \left[-a_{ij}u_i \left(\frac{u_t + \nu (t+1)^{\nu-1} u}{\nu} \exp (2(t+1)^\nu)\right)\right]_{x_i}
\]
\[+ \sum_{i,j=1}^{n} \left[-a_{ij}u_j \left(\frac{u_t + \nu (t+1)^{\nu-1} u}{\nu} \exp (2(t+1)^\nu)\right)\right]_{x_j} \tag{2.18}
\]
Next, we estimate from the above the left hand side of inequality (2.18) as
\[
(u_t - Lu)^2 \exp \left(2 (t + 1)^\nu \right) + \sqrt{\nu} (u_t - Lu) u \exp \left(2 (t + 1)^\nu \right)
\]
\[
\leq \frac{3}{2} (u_t - Lu)^2 \exp \left(2 (t + 1)^\nu \right) + \frac{\nu}{2} u^2 \exp \left(2 (t + 1)^\nu \right). \tag{2.19}
\]

Choose \(\nu_0\) such that, in addition to (2.10), (2.15) and (2.17)
\[
\nu_0 > 4. \tag{2.20}
\]
Then, \((\nu/2) u^2 < (\nu^2/8) u^2\) for \(\nu \geq \nu_0\). Hence, comparing (2.18) with (2.19), we obtain
\[
(u_t - Lu)^2 \exp \left(2 (t + 1)^\nu \right)
\]
\[
\geq \left[ \frac{\mu_1}{6} \sqrt{\nu} \left(\nabla u \right)^2 + \frac{\nu^2}{12} (t + 1)^{\nu-2} u^2 \right] \exp \left(2 (t + 1)^\nu \right) \tag{2.21}
\]
\[
+ \left[ \frac{2}{3} \left( \sum_{i,j=1}^n a_{ij} u_i u_j - \nu (t + 1)^{\nu-1} u^2 + \sqrt{\nu} u^2 \right) \right] \exp \left(2 (t + 1)^\nu \right),
\]
where the vector function \(U\) is such that
\[
U \mid_{\mathcal{S}} = 0. \tag{2.22}
\]
Condition (2.22) follows from the boundary condition in (2.3) and the lines 4–6 of (2.18).

**Step 6.** Integrate the pointwise Carleman estimate (2.21) over the domain \(Q_T\). Using Gauss’ formula and (2.22), we obtain for \(\nu \geq \nu_0\)
\[
\int_{Q_T} (u_t - Lu)^2 \exp \left(2 (t + 1)^\nu \right) \, dx \, dt
\]
\[
\geq \frac{\mu_1}{6} \sqrt{\nu} \int_{Q_T} (\nabla u)^2 \exp \left(2 (t + 1)^\nu \right) \, dx \, dt + \frac{\nu^2}{12} \int_{Q_T} u^2 \exp \left(2 (t + 1)^\nu \right) \, dx \, dt
\]
\[
+ \frac{2}{3} \exp \left(2 (T + 1)^\nu \right) \int_{\Omega} \left( \sum_{i,j=1}^n a_{ij} u_i u_j - \nu (T + 1)^{\nu-1} u^2 + \sqrt{\nu} u^2 \right) (x, T) \, dx
\]
\[
- \frac{2}{3} \nu^2 \int_{\Omega} \left( \sum_{i,j=1}^n a_{ij} u_i u_j \right) (x, 0) \, dx + \frac{2}{3} \nu t^2 \nu \int_{\Omega} \left( 1 - \frac{1}{2\sqrt{\nu}} \right) u^2 (x, 0) \, dx, \forall \nu \geq \nu_0.
\]
Step 7. Estimate from the below the third and fourth lines of the right hand side of (2.23) 
starting from the third line. Choose \( \nu_0 \) so large that, in addition to (2.10), (2.15) and (2.17), 
\[
\exp \left( 2 (T + 1)^{\nu} \right) \left( T + 1 \right)^{\nu - 1} < \exp \left( 3 (T + 1)^{\nu} \right), \quad \nu \geq \nu_0. 
\] 
(2.24)

By (1.3) 
\[
\left( \sum_{i,j=1}^{n} a_{i,j} u_{x_i} u_{x_j} \right) (x, T) \geq \mu_1 \left( \nabla u \right)^2 (x, T).
\]

Hence, the third line in the right hand side of (2.23) can be estimated as 
\[
\frac{2}{3} \exp \left( 2 (T + 1)^{\nu} \right) \int_{\Omega} \left( \sum_{i,j=1}^{n} a_{i,j} u_{x_i} u_{x_j} - \nu (T + 1)^{\nu - 1} u^2 + \sqrt{\nu} u^2 \right) (x, T) \, dx 
\geq - \exp \left( 3 (T + 1)^{\nu} \right) \| u \|_{H^1(\Omega)}.
\] 
(2.25)

We now work with the fourth line in the right hand side of (2.23). Obviously for \( \nu \geq \nu_0 > 1 \) 
\[
e^2 \nu \int_{\Omega} \left( 1 - \frac{1}{2 \sqrt{\nu}} \right) u^2 (x, 0) \, dx \geq 0.
\] 
(2.26)

Next, by (1.3) 
\[
- \frac{2}{3} e^2 \int_{\Omega} \left( \sum_{i,j=1}^{n} a_{i,j} u_{x_i} u_{x_j} \right) (x, 0) \, dx \geq - \mu_2 e^2 \| \nabla u (x, 0) \|^2.
\] 
(2.27)

Combining estimates (2.23) and (2.25)–(2.27), we obtain that for the choice of \( \nu_0 \) being so 
large that inequalities (2.10), (2.15), (2.17), (2.20) and (2.24) hold simultaneously, the target 
estimate (2.2) of this theorem is true for all \( \nu \geq \nu_0 \).

3. Stability estimate

For an arbitrary \( \tau \in (0, T) \), denote 
\[
H^{1,0} (Q_T) = \left\{ u : \| u \|_{H^{1,0} (Q_T)} = \left[ \int_{Q_T} \left( (\nabla u)^2 + u^2 \right) (x, t) \, dx \, dt \right]^{1/2} < \infty \right\}.
\]

Prior establishing our stability estimate, we prove lemma 3.1.

Lemma 3.1. Let \( \delta \in (0, 1) \) be a sufficiently small number and let the number \( k > 0 \). Choose 
a sufficiently large number \( \nu = \nu (\delta) \) such that 
\[
\exp \left( k (T + 1)^{\nu (\delta)} \right) = \frac{1}{\delta^4}.
\] 
(3.1)
Then for any $\tau \in (0, T)$ and for any number $y > 0$
\[
\lim_{\delta \to 0} \frac{\exp\left(-2 (\tau + 1)^{\nu(\delta)}\right)}{\delta^y} = \infty,
\]

\[
\lim_{\delta \to 0} \frac{\exp\left(-2 (\tau + 1)^{\nu(\delta)}\right)}{(\ln (\delta^{-1}))^{-\gamma}} = 0.
\]

(3.2) \hspace{1cm} (3.3)

**Remark 3.1.** It follows (3.2) and (3.3) that any stability estimate via $\exp (-2 (\tau + 1)^{\nu})$ is weaker than Hölder and stronger than the logarithmic stability estimate.

**Proof of lemma 3.1.** By (3.1) $k (T + 1)^{\nu(\delta)} = \ln (\delta^{-1})$. Hence, $(T + 1)^{\nu(\delta)} = \ln (\delta^{-1/k})$.

Hence,
\[
\nu(\delta) = \ln \left[ \left( \ln (\delta^{-1/k}) \right)^{1/\ln(T+1)} \right].
\]

(3.4)

Next,
\[
(\tau + 1)^{\nu(\delta)} = \exp \left( \ln (\tau + 1)^{\nu(\delta)} \right) = \exp (\nu(\delta) \ln (\tau + 1)).
\]

(3.5)

By (3.4)
\[
\nu(\delta) \ln (\tau + 1) = \ln \left[ \left( \ln (\delta^{-1/k}) \right)^{\ln(\tau+1)/\ln(T+1)} \right] = \ln \left[ \left( \ln (\delta^{-1/k}) \right)^{c} \right],
\]

(3.6)

\[
c = c(\tau, T) = \frac{\ln (\tau + 1)}{\ln (T + 1)} \in (0, 1).
\]

(3.7)

Using (3.5)–(3.7), we obtain
\[
(\tau + 1)^{\nu(\delta)} = \left( \ln (\delta^{-1/k}) \right)^{c} = \frac{1}{k^c} \left( \ln (\delta^{-1}) \right)^{c}.
\]

Hence,
\[
\exp \left(-2 (\tau + 1)^{\nu(\delta)}\right) = \exp \left[\frac{-2}{k^c} \left( \ln (\delta^{-1}) \right)^{c}\right].
\]

(3.8)

Using (3.8), we now prove (3.2). Indeed,
\[
\frac{\exp \left[-2 \left( \ln (\delta^{-1}) \right)^{c} / k^c\right]}{\delta^y} = \frac{\exp \left[-2 \left( \ln (\delta^{-1}) \right)^{c} / k^c\right]}{\exp (-y \ln (\delta^{-1}))}
\]

(3.9)
\[= \exp \left[ y \ln (\delta^{-1}) \left( 1 - \frac{2}{yk} (\ln (\delta^{-1}))^{c-1} \right) \right]. \]

Since by (3.7) \( c \in (0, 1) \), then
\[\lim_{\delta \to 0} \frac{2}{yk} (\ln (\delta^{-1}))^{c-1} = 0.\]

Hence,
\[\lim_{\delta \to 0} \left\{ \exp \left[ y \ln (\delta^{-1}) \left( 1 - \frac{2}{yk} (\ln (\delta^{-1}))^{c-1} \right) \right] \right\} = \infty.\]

This and (3.9) prove (3.2).

We now prove (3.3), which is equivalent with
\[\lim_{\delta \to 0} \frac{\exp \left[ -2 (\ln (\delta^{-1}))^{c} / k^c \right]}{\exp \left[ -y \ln (\ln (\delta^{-1})) \right]} = 0. \tag{3.10}\]

Next,
\[\frac{\exp \left[ -2 (\ln (\delta^{-1}))^{c} / k^c \right]}{\exp \left[ -y \ln (\ln (\delta^{-1})) \right]} = \exp \left[ -\frac{2}{k^c} (\ln (\delta^{-1}))^{c} + y \ln (\ln (\delta^{-1})) \right], \]
\[-\frac{2}{k^c} (\ln (\delta^{-1}))^{c} + y \ln (\ln (\delta^{-1})) = -\frac{2}{k^c} (\ln (\delta^{-1}))^{c} \left[ 1 - \frac{yk \ln (\ln (\delta^{-1}))}{2 (\ln (\delta^{-1}))^{c}} \right]. \tag{3.11}\]

Obviously
\[\lim_{\delta \to 0} \frac{\ln (\ln (\delta^{-1}))}{(\ln (\delta^{-1}))^{c}} = 0. \tag{3.12}\]

Thus, (3.10) follows from (3.11) and (3.12). \qed

**Theorem 3.1 (Stability estimate).** Assume that conditions (1.1)–(1.5) are in place. Suppose that two functions \( u_1, u_2 \in H^2(Q_T) \) are solutions of problem (1.7) and (1.8) with different data at \( \{ t = T \} \),
\[u_1(x, T) = g_1(x), \ u_2(x, T) = g_2(x), \ f(x) = g_1(x) - g_2(x). \tag{3.13}\]

Suppose that
\[\| f \|_{L^2(\Omega)} \leq \delta, \tag{3.14}\]
\[\| \nabla u_i(x, 0) \|_{L^2(\Omega)} \leq M, i = 1, 2, \tag{3.15}\]
where \( M > 0 \) is a number and the parameter \( \delta \) characterizes the level of noise in the data \( u(x, T) \). Denote \( \overline{C} > 0 \) be the number in (1.6) and \( \nu_0 \) be the parameter of
\textbf{Theorem 2.1.} Then there exist constants
\[ \nu_1 = \nu_1 \left( \mu_1, \mu_2, \max_{ij} \| a_{ij} \|_{C^1(\Omega)}, Q_T, C \right) \geq \nu_0 \] (3.16)
\[ C_1 = C_1 \left( \mu_1, \mu_2, \max_{ij} \| a_{ij} \|_{C^1(\Omega)}, Q_T, C \right) > 0, \] (3.17)
depending only on listed parameters such that if the number \( \delta_0 \) is so small that
\[ \ln \left( \frac{1}{\delta_0} \right) \geq \nu_1, \] (3.18)
then the following estimate holds for any \( \tau \in (0, T) \) and for all \( \delta \in (0, \delta_0) \)
\[ \| w \|_{H^1,0(Q_T)} \leq C_1 (M + 1) \exp \left[ -\frac{1}{\delta} \left( \ln \left( \frac{1}{\delta-1} \right) \right)^2 \right], \] (3.19)
where the constant \( C_1 \) is independent on \( M \) and the number \( c = c(\tau, T) \in (0, 1) \) is defined in (3.7).

Estimate (3.19) is between H"older and logarithmic stability estimates, see lemma 3.1 and remark 3.1. Everywhere below \( C_1 > 0 \) denotes different numbers depending on the same parameters as ones listed in (3.17).

\textbf{Proof.} It follows from (1.6)–(1.9) and (3.13) that
\[ |w_t - Lw| \leq C \left( |\nabla w| + |w| \right) \text{ a.e. in } Q_T, \] (3.20)
\[ w \mid_{\partial T} = 0, \] (3.21)
\[ w(x, T) = f(x). \] (3.22)
Let \( \nu \geq \nu_0 \). Square both sides of inequality (3.20), then multiply by \( \exp \left( (t + 1)^\nu \right) \) integrate over \( Q_T \) and then apply theorem 2.1 taking into account (3.21) and (3.22). We obtain
\[ 2C^2 \int_{Q_T} \left( |\nabla w|^2 + w^2 \right) \exp \left( (t + 1)^\nu \right) \, dx \, dt \]
\[ \geq \int_{Q_T} (w_t - Lw)^2 \exp \left( (t + 1)^\nu \right) \, dx \, dt \]
\[ \geq \frac{\mu_1}{6} \sqrt{\nu} \int_{Q_T} \left( |\nabla w|^2 + w^2 \right) \exp \left( (t + 1)^\nu \right) \, dx \, dt + \frac{\nu^2}{12} \int_{Q_T} w^2 \exp \left( (t + 1)^\nu \right) \, dx \, dt \]
\[ - \exp \left( (T + 1)^\nu \right) \| f \|^2_{L_2(\Omega)} - \mu_2 e^2 \| \nabla w (x, 0) \|^2_{L_2(\Omega)}, \forall \nu \geq \nu_0. \] (3.23)
Choose the number \( \nu_1 \) depending on the same parameters as ones listed in (3.16) such that \( \nu_1 \geq \nu_0 > 1 \) and satisfying the following two inequalities:
\[ \mu_1 \sqrt{\nu_1} \geq 24C^2, \quad \nu_1^2 \geq 48C^2. \] (3.24)
Then (3.14) and (3.23) imply that for all \( \nu \geq \nu_1 \)

\[
\delta^2 \exp \left( 3 (T + 1)^\nu \right) + \| \nabla w (x, 0) \|_2^2 \\
\geq \frac{\mu_1}{12} \sqrt{\nu} \int_{Q_r} (\nabla w)^2 \exp \left( 2 (t + 1)^\nu \right) \, dx \, dt + \frac{\nu^2}{24} \int_{Q_r} w^2 \exp \left( 2 (t + 1)^\nu \right) \, dx \, dt
\]

(3.25)

\[
\geq \frac{\mu_1}{12} \sqrt{\nu} \int_{Q_{r, r}} (\nabla w)^2 \exp \left( 2 (t + 1)^\nu \right) \, dx \, dt + \frac{\nu^2}{24} \int_{Q_{r, r}} w^2 \exp \left( 2 (t + 1)^\nu \right) \, dx \, dt
\]

\[
\geq \tilde{C}_1 \exp \left( 2 (\tau + 1)^\nu \right) \| w \|_{H^{1/2} (Q_{r, r})}^2
\]

(3.26)

Hence,

\[
\| w \|_{H^{1/2} (Q_{r, r})}^2 \leq C_1 \delta^2 \exp \left( 3 (T + 1)^\nu \right) + C_1 \| \nabla w (x, 0) \|_2^2 \exp \left( -2 (\tau + 1)^\nu \right)
\]

(3.27)

where the new value of \( C_1 \) is \( C_1 = 1/\tilde{C}_1 \) : see (3.24), (3.26) and also recall that \( C_1 > 0 \) denotes different positive constants depending on parameters listed in (3.17).

Choose \( \nu = \nu (\delta) \) such that (3.1) would be satisfied with \( k = 3 \), i.e.

\[
\exp \left( 3 (T + 1)^\nu \right) = \frac{1}{\delta}
\]

(3.28)

Hence,

\[
\nu = \nu (\delta) = \ln \left[ \left( \ln \left( \delta^{-1/3} \right) \right)^{1/\ln(T+1)} \right].
\]

(3.29)

The choice (3.29) is possible since (3.18) holds and \( \delta \in (0, \delta_0) \). Hence, by (3.8)

\[
\exp \left( -2 (\tau + 1)^\nu (\delta) \right) = \exp \left[ -\frac{2}{3} \left( \ln \left( \delta^{-1} \right) \right)^c \right].
\]

(3.30)

Hence, (3.15), (3.27), (3.28) and (3.30) imply that

\[
\| w \|_{H^{1/2} (Q_{r, r})} \leq C_1 \delta + 2 C_1 M^2 \exp \left[ -\frac{2}{3} \left( \ln \left( \delta^{-1} \right) \right)^c \right].
\]

(3.31)

The target estimate (3.19) of this theorem obviously follows from (3.2) and (3.31).

\( \square \)

4. The quasi-reversibility method for the linear case

We assume in this section that the function \( F (\nabla u, u, x, t) \) in (1.7) is linear with respect to the function \( u \) and its first derivatives,
\[ F(\nabla u, u, x, t) = Bu + p(x, t) = \sum_{j=1}^{n} b_j(x,t) u_{x_j} + c(x, t) u + p(x, t), \quad (4.1) \]

\[ Bu = \sum_{j=1}^{n} b_j(x,t) u_{x_j} + c(x, t) u, \quad (4.2) \]

where functions \( b_j, c, p \in C(\mathcal{Q}_T) \). Then (1.7)–(1.9) become
\[ u_t = Lu + Bu + p(x, t) , (x, t) \in Q_T, \quad (4.3) \]
\[ u \mid_{S_T} = 0, \quad (4.4) \]
\[ u(x, T) = g(x), x \in \Omega. \quad (4.5) \]

Assuming that \( g(x) = 0 \) for \( x \in \partial \Omega \) and
\[ g \in H^2(\Omega), \quad (4.6) \]
consider the function
\[ v(x, t) = u(x, t) - g(x). \quad (4.7) \]

Then (4.1)–(4.7) lead to
\[ v_t = Lv + Bv + q(x, t), (x, t) \in Q_T, \quad (4.8) \]
\[ v \mid_{S_T} = 0, \quad (4.9) \]
\[ v(x, T) = 0, x \in \Omega, \quad (4.10) \]
\[ q(x, t) = (Lg + Bg + p)(x, t) \in L^2(Q_T). \quad (4.11) \]

We introduce the subspace \( H^0_0(Q_T) \) of the space \( H^2(Q_T) \) as
\[ H^0_0(Q_T) = \{ w \in H^2(Q_T) : w \mid_{S_T} = 0, w(x, T) = 0 \}. \]

The QRM for problem (4.8)–(4.11) amounts to the minimization of the following functional
\[ J_\alpha(v) = \int_{Q_T} (v_t - Lv - Bv - q)^2 \, dx dt + \alpha \| v \|_{H^2(Q_T)}^2, \quad (4.12) \]
where \( \alpha \in (0, 1) \) is the regularization parameter. We arrive at the following problem:

**Minimization problem 1.** Minimize the functional \( J_\alpha(v) \) on the space \( H^0_0(Q_T) \).

**Theorem 4.1.** Assume that conditions (1.1)–(1.5) and (4.11) hold. Then there exists unique minimizer \( v_{\text{min}} \in H^0_0(Q_T) \) of the functional \( J_\alpha(v) \).

**Proof.** Let \([,] \) denotes the scalar product in \( H^2(Q_T) \). By the variational principle, any minimizer \( v_{\text{min}} \in H^0_0(Q_T) \), if it exists, satisfies the following integral identity
\[ \int_{Q_T} (\partial_t v_{\text{min}} - Lv_{\text{min}} - Bv_{\text{min}})(h_t - Lh - Bh) \, dx dt + \alpha \| v_{\text{min}} \|_{H^2(Q_T)}^2 \quad (4.13) \]
Since
\[ \int_{Q_T} (v_t - Lv - Bv)^2 \, dx \, dt + \alpha [v, v]^2 \geq 2 \alpha \|v\|^2_{H^1_0(Q_T)}, \forall v \in H^2_0(Q_T), \]
then the equality
\[ \{v, h\} = \int_{Q_T} (v_t - Lv - Bv) (h_t - Lh - Bh) \, dx \, dt + \alpha [v, h], \forall v, h \in H^2_0(Q_T) \]
defines a new scalar product in the space $H^2_0(Q_T)$. We rewrite integral identity (4.13) as
\[ \{v_{\text{min}}, h\} = \int_{Q_T} q (h_t - Lh - Bh) \, dx \, dt. \] (4.14)
Next, by the Cauchy–Schwarz inequality
\[ \int_{Q_T} q (h_t - Lh - Bh) \, dx \, dt \leq \|q\|_{L_2(Q_T)} \|h_t - Lh - Bh\|_{L_2(Q_T)} \leq C_1 \|q\|_{L_2(Q_T)} \|h\|_{H^1_0(Q_T)}. \]
Hence, by Riesz theorem, there exists a unique function $s = s(q) \in H^2_0(Q_T)$ such that
\[ \int_{Q_T} q (h_t - Lh - Bh) \, dx \, dt = \{s, h\}, \forall h \in H^2_0(Q_T). \]
Comparing this with (4.14), we obtain
\[ \{v_{\text{min}}, h\} = \{s, h\}, \forall h \in H^2_0(Q_T). \]
Therefore, there exists unique minimizer $v_{\text{min}} = s \in H^2_0(Q_T)$ of the functional $J_0(v)$. □

In the regularization theory, the function $v_{\text{min}}$ is called the regularized solution of problem (4.8)–(4.11) [3, 29]. The next step after theorem 4.1 is to prove convergence of regularized solutions to the exact one when the noise in the data tends to zero. While we have used only Riesz theorem to prove existence and uniqueness of the minimizer, the convergence result requires the Carleman estimate of theorem 2.1.

To prove the convergence result, we need to introduce noise in the data $g(x)$ in (1.9). By (4.6) we assume that the noise in $g(x)$ is small in the norm of the space $H^2(\Omega)$. Following one of the Tikhonov’s concept of the regularization [3, 29], we assume that there exists exact noiseless data $g^* \in H^2(\Omega)$ for problem (4.3)–(4.5). Let
\[ \|g - g^*\|_{H^1(\Omega)} < \sigma, \] (4.15)
where $\sigma > 0$ is a small number characterizing the noise level in the data. Then (4.1), (4.2), (4.11) and (4.15) lead to
\[\|q - q^*\|_{L^2(Q_T)} = \|L(g - g^*) + B(g - g^*)\|_{L^2(Q_T)} \leq \|L(g - g^*)\|_{L^2(Q_T)} + \|B(g - g^*)\|_{L^2(Q_T)} \leq C_1\sigma,\]

where by (4.11) \(q^*(x, t) = (Lg^* + Bg^* + p)(x, t) \in L_2(Q_T)\). Hence, it is convenient for derivations below in this section to introduce \(\delta = C_1\sigma\) as the noise level and to assume that \(\|q - q^*\|_{L^2(Q_T)} < \delta\). 

(4.16)

**Remark 4.1.** We observe that going from the noisy function \(g(x)\) to the function \(q(x, t)\) via (4.11) and still keeping a reasonably small amount of noise \(\delta\) in (4.16) is a difficult computational procedure for the spatial dimension \(n > 2\). Indeed, by (4.11), one needs to obtain good approximations for the first and second derivatives of the function \(g(x)\). It seems that one might be able to adapt some regularization methods for that procedure. However, we are not discussing this issue here since it is outside of the scope of the current publication.

Again, let \(v_{\text{min}} \in H_0^2(Q_T)\) be the minimizer of the functional (4.12), the existence and uniqueness of which were established by theorem 4.1. Consider the difference

\[w_\delta = v_{\text{min}} - v^*.\]

Theorem 4.2 estimates the function \(w\) via \(\delta\). It follows from the regularization theory that we need to assume a certain dependence \(\alpha = \alpha(\delta)\) of the regularization parameter on the noise level \(\delta\).

**Theorem 4.2 (Convergence rate).** Assume that conditions (1.1)–(1.6) and (4.16) are in place. Let \(\nu_0 > 1\) be the number of theorem 2.1 and let \(\alpha = \alpha(\delta) = \delta^2\). Then there exists a number

\[\nu_2 = \nu_2\left(\mu_1, \mu_2, \max_{ij} \|a_{ij}\|_C(Q_T)\right) \geq \nu_0 > 1, \]

depending only on listed parameters such that if \(\delta \in (0, \delta_0)\) and \(\delta_0 \in (0, 1)\) is so small that

\[\ln \left[\left(\ln \left(\delta_0^{-1/2}\right)\right)^{1/\ln(T + 1)}\right] \geq \nu_2,\]

\(\delta_0 \leq \exp \left[-\frac{2}{\tau^c} \left(\ln \left(\delta_0^{-1}\right)\right)^c\right],\)

then the following convergence estimate of the QRM holds for every \(\tau \in (0, T):\)

\[\|w_\delta\|_{H^3(Q_T)} \leq C_1 \left(1 + \|v^*\|_{H^2(Q_T)}\right) \exp \left[-\frac{1}{2^c} \left(\ln \left(\delta^{-1}\right)\right)^c\right],\]

where the function \(w_\delta\) is defined in (4.17), the number \(c = c(\tau, T) \in (0, 1)\) is defined in (3.7).

As to (4.20), see (3.2) in lemma 3.1.

**Proof of theorem 4.2.** The function \(v^* \in H_0^2(Q_T)\) satisfies the following integral identity

\[\int_{Q_T} (v_t^* - L\nu^* - B\nu^*) \left(h_t - Lh - Bh\right) \, dx \, dt + \alpha \left[v^*, h\right]\]

(4.22)
\[
= \int q^\ast (h_t - Lh - Bh) \, dx \, dt + \alpha \|v^\ast\|^2_{H^1(\Omega)} , \forall h \in H_0^2 (\Omega_T).
\]

Subtracting (4.22) from (4.13) and using (4.17), we obtain
\[
\int_{\Omega_T} (w_{\delta t} - Lw_{\delta} - Bw_{\delta}) (h_t - Lh - Bh) \, dx \, dt + \alpha \|w_{\delta}\|^2_{H^1(\Omega_T)} \tag{4.23}
\]
\[
= \int_{\Omega_T} (q - q^\ast) (h_t - Lh - Bh) \, dx \, dt - \alpha \|v^\ast\|^2_{H^1(\Omega_T)} , \forall h \in H_0^{2,1} (\Omega_T).
\]

Set in (4.23) \(h = w_{\delta t}\). Using Cauchy–Schwarz inequality, we obtain
\[
\int_{\Omega_T} (w_{\delta t} - Lw_{\delta} - Bw_{\delta})^2 \, dx \, dt + \alpha \|w_{\delta}\|^2_{H^1(\Omega_T)} \tag{4.24}
\]
\[
\leq \frac{1}{2} \|q - q^\ast\|^2_{L_2(\Omega_T)} + \frac{1}{2} \int_{\Omega_T} (w_{\delta t} - Lw_{\delta} - Bw_{\delta})^2 \, dx \, dt + \frac{\alpha}{2} \|v^\ast\|^2_{H^1(\Omega_T)}
\]
\[
+ \frac{\alpha}{2} \|w_{\delta}\|^2_{H^1(\Omega_T)}.
\]

Hence, combining (4.24) with (4.16), we obtain
\[
\int_{\Omega_T} (w_{\delta t} - Lw_{\delta} - Bw_{\delta})^2 \, dx \, dt + \alpha \|w_{\delta}\|^2_{H^1(\Omega_T)} \leq \delta^2 + \alpha \|v^\ast\|^2_{H^1(\Omega_T)} \tag{4.25}
\]

Since \(\alpha = \alpha (\delta) = \delta^2\), then (4.25) implies that \(\|w_{\delta}\|^2_{H^1(\Omega_T)} \leq 1 + \|v^\ast\|^2_{H^1(\Omega_T)}\). Hence, trace theorem leads to
\[
\|\nabla w_{\delta} (x, 0)\|_{L_2(\Omega)} \leq C \left(1 + \|v^\ast\|^2_{H^1(\Omega_T)}\right). \tag{4.26}
\]

To proceed further, we apply to (4.25) the Carleman estimate of theorem 2.1. Let the number \(\nu_2\) satisfies (4.18) and let \(\nu \geq \nu_2\). We have
\[
\int_{\Omega_T} (w_{\delta t} - Lw_{\delta} - Bw_{\delta})^2 \, dx \, dt
\]
\[
= \int_{\Omega_T} \exp \left(2 (t + 1)^\nu\right) \exp \left(-2 (t + 1)^\nu\right) (w_{\delta t} - Lw_{\delta} - Bw_{\delta})^2 \, dx \, dt
\]
\[
\geq \exp \left(-2 (T + 1)^\nu\right) \int_{\Omega_T} (w_{\delta t} - Lw_{\delta} - Bw_{\delta})^2 \exp \left(2 (t + 1)^\nu\right) \, dx \, dt.
\]

Hence, using (4.25), we obtain
\[
\int_{\Omega_T} (w_{\delta t} - Lw_{\delta} - Bw_{\delta})^2 \exp \left(2 (t + 1)^\nu\right) \, dx \, dt \tag{4.27}
\]
\[ \leq \exp\left(2(T + 1)^\nu\right) \delta^2 \left(1 + \|v^*\|^2_{H^1(Q_T)}\right). \]

We have

\[ \int_{Q_T} (w_{st} - Lw_\delta - Bw_\delta)^2 \exp\left(2(t + 1)^\nu\right) \, dx \, dt \]

\[ \geq \int_{Q_T} (w_{st} - Lw_\delta)^2 \exp\left(2(t + 1)^\nu\right) \, dx \, dt \]

\[ -C_1 \int_{Q_T} \left((\nabla w_\delta)^2 + w_\delta^2\right)^2 \exp\left(2(t + 1)^\nu\right) \, dx \, dt. \]

Next, since \( w_\delta(x, T) = 0 \), then by (2.2)

\[ \int_{Q_T} (w_{st} - Lw_\delta)^2 \exp\left(2(t + 1)^\nu\right) \, dx \, dt \]

\[ -C_1 \int_{Q_T} \left((\nabla w_\delta)^2 + w_\delta^2\right)^2 \exp\left(2(t + 1)^\nu\right) \, dx \, dt \]

\[ \geq \frac{\mu_1}{4} \int_{Q_T} (\nabla w_\delta)^2 \exp\left(2(t + 1)^\nu\right) \, dx \, dt + \frac{\nu^2}{8} \int_{Q_T} w_\delta^2 \exp\left(2(t + 1)^\nu\right) \, dx \, dt \]

\[ -C_1 \int_{Q_T} \left((\nabla w_\delta)^2 + w_\delta^2\right)^2 \exp\left(2(t + 1)^\nu\right) \, dx \, dt - \mu_2 e^2 \|\nabla w_\delta(x, 0)\|^2_{L^2(\Omega)}. \]

Choose \( \nu_2 \geq \nu_0 > 1 \) such that \( \mu_1 \sqrt{\nu_2} \geq 8C_1 \) and \( \nu_2 \geq 4\sqrt{C_1}. \) Then, using (4.26) and (4.29), we obtain

\[ \int_{Q_T} (w_{st} - Lw_\delta)^2 \exp\left(2(t + 1)^\nu\right) \, dx \, dt \]

\[ -C_1 \int_{Q_T} \left((\nabla w_\delta)^2 + w_\delta^2\right)^2 \exp\left(2(t + 1)^\nu\right) \, dx \, dt \]

\[ \geq \frac{\mu_1}{8} \int_{Q_T} (\nabla w_\delta)^2 \exp\left(2(t + 1)^\nu\right) \, dx \, dt + \frac{\nu^2}{16} \int_{Q_T} w_\delta^2 \exp\left(2(t + 1)^\nu\right) \, dx \, dt \]

\[ -C_1 \mu_2 e^2 \left(1 + \|v^*\|^2_{H^1(Q_T)}\right) \]

\[ \geq C_1 \exp\left(2(T + 1)^\nu\right) ||w_\delta||^2_{H^1(\Omega)} - C_1 \mu_2 e^2 \left(1 + \|v^*\|^2_{H^1(Q_T)}\right), \forall \nu \geq \nu_2. \]
Hence, (4.27)–(4.30) imply that
\[
\exp \left(2 (T + 1)^\nu \right) \delta^2 \left(1 + \|v^*\|_{H^2(Q_T)}^2 \right) + C \mu_2 e^2 \left(1 + \|v^*\|_{H^2(Q_T)}^2 \right) \\
\geq C_1 \exp \left(2 (\tau + 1)^\nu \right) \|w^*\|_{H^2(Q_T)}, \forall \nu \geq \nu_2.
\]
Dividing this by \( \exp \left(2 (\tau + 1)^\nu \right) \), we obtain for all \( \nu \geq \nu_2 \)
\[
\|w^*\|_{H^2(Q_T)} \leq C_1^{-1} \left(1 + \|v^*\|_{H^2(Q_T)}^2 \right) \exp \left(2 (T + 1)^\nu \right) \delta^2 \] (4.31)
\[+ C C_1^{-1} \mu_2 e^2 \exp \left(-2 (\tau + 1)^\nu \right) \left(1 + \|v^*\|_{H^2(Q_T)}^2 \right).\]
Choose \( \nu = \nu(\delta) \) such that (3.1) would be satisfied with \( k = 2 \), i.e.
\[
\exp \left(2 (T + 1)^{\nu(\delta)} \right) = \frac{1}{\delta}.
\] (4.32)
Hence,
\[
\nu = \nu(\delta) = \ln \left( \ln \left( \delta^{-1/2} \right)^{1/\ln(T + 1)} \right).
\] (4.33)
The choice (4.32) and (4.33) is possible since (4.19) holds and \( \delta \in (0, \delta_0) \). It follows from (3.1), (3.8) and (4.32) that
\[
\exp \left(-2 (\tau + 1)^{\nu(\delta)} \right) = \exp \left[ -\frac{2}{C_2} \left( \ln \left( \delta^{-1} \right) \right)^{\epsilon} \right].
\] (4.34)
Since in our derivations \( C > 0 \) and \( C_1 > 0 \) denote different constants depending on parameters listed in (2.1) and (3.17) respectively, then we can regard the number \( \sqrt{\max \left( C_1^{-1}, C_1^{-1} \mu_2 e^2 \right)} \) as \( C_1/\sqrt{2} \) with a new value of \( C_1 \). Hence, (4.31)–(4.34) imply that
\[
\|w^*\|_{H^2(Q_T)} \leq C_2^2 \left(1 + \|v^*\|_{H^2(Q_T)}^2 \right) \left\{\delta + \exp \left[ -\frac{2}{C_2^2} \left( \ln \left( \delta^{-1} \right) \right)^{\epsilon} \right]\right\}. \] (4.35)
Note that inequality (4.20) remains true if \( \delta_0 \) is replaced with \( \delta \in (0, \delta_0) \). Hence, using (4.35), we obtain
\[
\|w^*\|_{H^2(Q_T)} \leq C_2^2 \left(1 + \|v^*\|_{H^2(Q_T)}^2 \right) \exp \left[ -\frac{2}{C_2^2} \left( \ln \left( \delta^{-1} \right) \right)^{\epsilon} \right],
\]
which implies (4.21).
\[\square\]

5. The global strict convexity

While the linear case (4.1) was studied in section 4, in this section we consider the quasilinear case.
5.1. The Tikhonov-like functional with the CWF in it

Recall that by (1.9) the function \( g(x) \) is defined as \( g(x) = u(x, T) \). Since we need below the function \( (Lg)(x, t) \) to be bounded in \( \overline{Q_T} \), then we assume that 
\[
    g \in C^2 (\Omega) .
\] (5.1)

Just like in (4.7), we consider the function \( v(x, t) = u(x, t) - g(x) \). Then, assuming (5.1), we obtain instead of (1.7)–(1.9):
\[
    v_t = Lv + G ( \nabla v, v, g, x, t ) , (x, t) \in Q_T , \quad v |_{S_T} = 0 , \quad v(x, T) = 0, x \in \Omega , \quad v(x, t) = 0 , x \in \Omega , \quad G ( \nabla v, v, g, x, t ) = Lg + F ( \nabla v + \nabla g, v + g, x, t ) .
\] (5.5)

In our derivations below \( v_t, Lv \) as well as arguments of the function \( F ( \nabla v + \nabla g, v + g, x, t ) \) must be uniformly bounded for all \( (x, t) \in \overline{Q_T} \). Hence, similarly with [1, 17], we now need to impose a higher smoothness than just \( v \in H^2_0 (Q_T) \) as in section 4. Consider an integer \( k > [(n + 1) / 2] + 2 \), where \( [(n + 1) / 2] \) denotes the maximal integer which does not exceed \( (n + 1) / 2 \). Then embedding theorem implies that \( H^k (Q_T) \subset C^2 (\overline{Q_T}) \) and 
\[
    \| f \|_{C^2 (\overline{Q_T})} \leq E \| f \|_{H^k (Q_T)} , \forall f \in H^k (Q_T) .
\] (5.6)

where the number \( E = E (Q_T) > 0 \) depends only on the domain \( Q_T \). Define the subspace \( H^0_k (Q_T) \subset H^k (Q_T) \) as
\[
    H^0_k (Q_T) = \{ v \in H^k (Q_T) : v |_{S_T} = 0 , v(x, T) = 0 \} .
\]

Let \( R > 0 \) be an arbitrary number. We consider the ball \( B(R) \) in the space \( H^0_k (Q_T) \),
\[
    B(R) = \{ v \in H^0_k (Q_T) : \| v \|_{H^k (Q_T)} < R \} .
\] (5.7)

Hence, by (5.6)
\[
    B(R) \subset C^2 (\overline{Q_T}) ,
\] (5.8)
\[
    \| v \|_{C^2 (\overline{Q_T})} \leq E_k , \forall v \in B(R) ,
\] (5.9)

where the number \( E_k = E_k (Q_T, R) = \text{const.} > 0 \) depends only on listed parameters.

We want to find an approximate solution of problem (5.2)–(5.5) in the closed ball \( B(R) \).

To do this, we select a number \( \tau \in (0, T) \) and minimize the following weighted Tikhonov-like functional
\[
    I_{\alpha, \nu} (v) = \exp (-2 (\tau + 1)^\nu) \int_{Q_T} (v_t - Lv - G ( \nabla v, v, x, t ) )^2 \exp (2 (\tau + 1)^\nu) \, dx \, dt + \alpha \| v \|_{H^k (Q_T)}^2 , v \in B(R) .
\] (5.10)

The multiplier \( \exp (-2 (\tau + 1)^\nu) \) in (5.10) introduced to balance two terms in the right hand side of (5.10). Indeed, the regularization parameter \( \alpha \in (0, 1) \) and
\[
\min_{\tau \in [0,T]} \left\{ \exp (2 (t + 1)^\nu) \exp (-2 (\tau + 1)^\nu) \right\} = 1. 
\] (5.11)

also, see (5.15).

**Minimization problem 2.** Minimize the functional \( I_{\alpha, \lambda, \nu} (v) \) on the ball \( \overline{B}(R) \).

5.2. **Theorems about the functional \( I_{\alpha, \nu} (v) \)**

The central theorem of this section is theorem 5.1.

**Theorem 5.1 (Global strict convexity).** Assume that conditions (1.1)–(1.6) and (5.5) hold. Then the functional \( I_{\alpha, \nu} (v) \) has the Fréchet derivative \( I'_{\alpha, \nu} (v) \in H^1_Q(\tilde{Q}_T) \) at every point \( v \in H^1_Q(\tilde{Q}_T) \) and for all values of parameters \( \alpha, \nu \geq 0 \). Let \( \nu_0 > 1 \) be the number of theorem 2.1. Then there exists numbers

\[
\nu_3 = \nu_3 \left( \mu_1, \mu_2, \max_{ij} |a_{ij}| c^1(\overline{\Omega}), \|g\|_{C^1(\overline{\Omega})}, \|Q_T, R, C, \tau\| \right) \geq \nu_0 > 1, \quad (5.12)
\]

\[
C_2 = C_2 (\mu_1, \mu_2, \max_{ij} |a_{ij}| c^1(\overline{\Omega}), \|g\|_{C^1(\overline{\Omega})}, \|Q_T, R, C\|) > 0 \quad (5.13)
\]

depending only on listed parameters such that \( 2C_2 \mu_2 e^2 \exp (-2 (\tau + 1)^\nu) \in (0, 1) \) for \( \nu \geq \nu_3 \) and \( \alpha \in \left[ 2C_2 \mu_2 e^2 \exp (-2 (\tau + 1)^\nu), 1 \right), \quad (5.14) \)

then the functional \( I_{\alpha, \nu} (v) \) is strictly convex on \( \overline{B}(R) \). More precisely, for every \( \tau \in (0, T) \) the following strict convexity estimate holds

\[
I_{\alpha, \nu} (v_2) - I_{\alpha, \nu} (v_1) - I'_{\alpha, \nu} (v_1) (v_2 - v_1)
\geq C_2 \|v_2 - v_1\|_{H^\mu(Q_T)}^2 + \frac{\alpha}{2} \|v_2 - v_1\|_{H^\nu(Q_T)}^2, \quad \forall v_1, v_2 \in \overline{B}(R). \quad (5.15)
\]

Everywhere below \( C_2 > 0 \) denotes different positive constants depending on the same parameters as those listed in (5.13).

**Remark 5.1.** The presence of the term \( C_2 \|v_2 - v_1\|_{H^\mu(Q_T)}^2 \) in the right hand side of (5.15) indicates that the convergence of a gradient-like method of the minimization of functional \( I_{\alpha, \nu} (v) \) is likely faster in the space \( H^{1, \mu} (Q_T) \) than in the space \( H^\nu (Q_T) \).

**Theorem 5.2.** The Fréchet derivative \( I'_{\alpha, \nu} (v) \) of the functional \( I_{\alpha, \nu} (v) \) is Lipschitz continuous on \( B(2R) \) for all values of parameters \( \alpha, \lambda, \nu \geq 0 \). In other words, there exists a number

\[
D = D \left( \mu_1, \mu_2, \max_{ij} |a_{ij}| c^1(\overline{\Omega}), \|g\|_{H^\mu(\Omega)}, \|Q_T, R, C, \lambda, \nu, \alpha\| \right) > 0
\]

depending only on listed parameters such that

\[
\|I'_{\alpha, \nu} (v_2) - I'_{\alpha, \nu} (v_1)\|_{H^\nu(Q_T)} \leq D \|v_2 - v_1\|_{H^\nu(Q_T)}, \quad \forall v_1, v_2 \in B(2R). \quad (5.16)
\]
Furthermore, let \( \nu_3 \) be the number of theorem 5.1. Then for every pair \( \alpha > 0, \nu \geq \nu_3 \) there exists unique minimizer \( v_{\min} \in \overline{B(R)} \) of the functional \( I_{\alpha,\nu} (v) \) on the closed ball \( \overline{B(R)} \) and the following inequality holds
\[
I'_{\alpha,\nu}(v_{\min}) (v_{\min} - w) \leq 0, \quad \forall w \in \overline{B(R)}. \tag{5.17}
\]

Let \( P_{\Pi} : H_0^1(Q_T) \to \overline{B(R)} \) be the orthogonal projection operator mapping the space \( H_0^1(Q_T) \) onto the closed ball \( \overline{B(R)} \). Let \( v_0 \in B(R) \) be an arbitrary point of \( B(R) \). Let the number \( \gamma \in (0, 1) \). Consider the sequence of the gradient projection method,
\[
v_n = P_{\Pi} (v_{n-1} - \gamma I'_{\alpha,\nu}(v_{n-1})) , \quad n = 1, 2, \ldots \tag{5.18}
\]

**Theorem 5.3.** Let \( \nu_3 \) be the number of theorem 5.1. Choose the number \( \nu \geq \nu_3 \). Let \( v_{\min} \in \overline{B(R)} \) be the unique minimizer of the functional \( I_{\alpha,\nu} (v) \) on the set \( \overline{B(R)} \) (theorem 5.2). Then there exists a sufficiently small number
\[
\gamma_0 = \gamma_0 \left( \mu_1, \mu_2, \max_{ij} \|a_{ij}\|_{C(\overline{\Omega}_T)}, \|g\|_{C(\overline{\Omega}_T)} , Q_T, R, C, \nu, \alpha \right) \in (0, 1)
\]
such that for every \( \gamma \in (0, \gamma_0) \) there exists a number \( \theta = \theta (\gamma) \in (0, 1) \) such that
\[
\|v_n - v_{\min}\|_{H^1(Q_T)} \leq \theta^n \|v_{\min} - v_0\|_{H^0(Q_T)}. \tag{5.19}
\]

Consider now the case of noise in the data. Similarly with section 4, we use again one of the Tikhonov’s concept of the regularization. More precisely, using (5.1), we assume that there exists exact noiseless data \( g^* \in C^2(\overline{\Omega}) \) in (5.5) and, respectively, there exists exact solution \( v^* \in B(R) \) of problem (5.2)–(5.5), where \( g \) is replaced with \( g^* \). Let \( \delta \in (0, 1) \) be the level of noise in the data \( g (x) \), i.e.
\[
\|g - g^*\|_{C^2(\overline{\Omega})} < \delta. \tag{5.20}
\]

In theorem 5.4 we estimate the accuracy of the minimizer, i.e. the norm \( \|v_{\min} - v^*\|_{H^0(Q_{\tau_T})} \) for any \( \tau \in (0, T) \). In turn, this estimate, combined with (5.19), provides the convergence rate of the sequence (5.18) to the exact solution. Note that since \( \delta \in (0, 1) \), then by (5.20), we replace below dependencies of the above numbers \( \nu_3, C_2, D, \gamma_0 \) on \( \|g\|_{C^2(\overline{\Omega})} \) with their dependencies on \( \|g^*\|_{C^2(\overline{\Omega})} \).

**Theorem 5.4 (Estimates of the accuracy and the convergence rate).** Assume that the exact solution of problem (5.2)–(5.6) \( v^* \in B(R) \) and that (5.20) holds. Let \( \nu_3 > 1 \) be the number of theorem 5.1. Select an arbitrary number \( \tau \in (0, T) \). For any \( \delta \in (0, 1) \) set the number \( \nu = \nu (\delta) \) be the same as in (4.33). Let the number
\[
\delta_0 = \delta_0 \left( \mu_1, \mu_2, \max_{ij} \|a_{ij}\|_{C(\overline{\Omega}_T)}, \|g^*\|_{C^2(\overline{\Omega})} , Q_T, R, C \right) > 0
\]
be so small that
\[
\nu (\delta_0) > \nu_3 \quad \text{and} \quad 2C_2 \mu_2 e^2 \exp \left( -2 (\tau + 1)^{\nu_3 (\delta_0)} \right) \in (0, 1). \tag{5.21}
\]
Let $\delta \in (0, \delta_0)$ and let $v_{\text{min}} \in \overline{B}(R)$ be the unique minimizer of the functional $I_{\alpha, \nu}(v)$ on the set $\overline{B}(R)$ (theorem 5.2). Let $\gamma_0$ be the number defined in theorem 5.3. Let $\gamma \in (0, \gamma_0)$ and $\theta = \theta(\gamma) \in (0, 1)$ be also the numbers of theorem 5.3. Choose the regularization parameter $\alpha$ as

$$
\alpha = \alpha(\delta) = 2C_2\mu_2 e^{2 \exp \left( -2 (\tau + 1)^{\nu(\delta)} \right)}.
$$

(5.22)

Then the following accuracy and convergence estimates hold

$$
\|v^* - v_{\text{min}}\|_{H^{s}(Q_T)} \leq C_2 \exp \left[ -\frac{1}{2c}(\ln (\delta^{-1}))^c \right],
$$

(5.23)

$$
\|v^* - v_n\|_{H^{s}(Q_T)} \leq C_2 \exp \left[ -\frac{1}{2c}(\ln (\delta^{-1}))^c \right] + \theta^n \|v_{\text{min}} - v_0\|_{H^s(Q_T)},
$$

(5.24)

where the number $c = c(\tau, T) \in (0, 1)$ is defined in (3.7).

**Remark 5.2.** According to section 1, since $R > 0$ is an arbitrary number and since the starting point $v_0 \in \overline{B}(R)$ of the gradient projection method is an arbitrary point of $\overline{B}(R)$, then theorem 5.4 implies the global convergence to the exact solution of the gradient projection method of the minimization of the functional $I_{\alpha, \nu}(v)$.

In this paragraph, we temporarily assume that theorem 5.1 is proved. Then the proof of the Lipschitz continuity (5.16) of the Fréchet derivative $I'_{\alpha, \nu}$ is very similar to the proof of theorem 3.1 of [1]. The rest of theorem 5.2 follows from (5.16) and lemma 2.1 of [1]. Given theorems 5.1 and 5.3 follows from theorem 3.3 of [1].

Therefore, we prove in section 6 only theorems 5.1 and 5.4. Below, if we say that a vector function belongs to a certain Banach space, then this means that each of its components belongs to that space. The norm of that vector function in that space is defined as the square root of the sum of squared norms of its components.

### 6. Proofs of theorems 5.1 and 5.4

#### 6.1. Proof of theorem 5.1

Consider two arbitrary functions $v_1, v_2 \in \overline{B}(R)$. Denote $h = v_2 - v_1$. Then

$$
h \in \overline{B}(2R).
$$

(6.1)

Hence, by (5.9)

$$
h \in C^2(\overline{Q}_T), \|h\|_{C^2(\overline{Q}_T)} \leq E_{2R}.
$$

(6.2)

Using the multidimensional analog of Taylor formula [31], (1.6), (5.5), (5.9) and (6.2), we obtain

$$
G(\nabla v_2, v_2, x, t) = G(\nabla v_1 + \nabla h, v_1 + h, x, t)
$$

(6.3)
where the $n - D$ vector function $G_1 (x, t) \in C (\Theta_T)$, the function $G_2 (x, t) \in C (\Theta_T)$ and the function $G_3 (x, t, \nabla h, h)$ is such that
\[
|G_3 (x, t, \nabla h, h)| \leq C_2 \left( (\nabla h)^2 + h^2 \right) (x, t), \forall h \in B (2R), \forall (x, t) \in \overline{\Theta}_T. \tag{6.4}
\]
In addition,
\[
\|G_1 (x, t)\|_{C(\Theta_T)}, \|G_2 (x, t)\|_{C(\Theta_T)} \leq C_2. \tag{6.5}
\]
Hence,
\[
v_{2t} - Lv_2 - G (\nabla v_2, v_2, x, t)
= [h_t - Lh - G_1 (x, t) \nabla h - G_2 (x, t) h - G_3 (x, t, \nabla h, h)]
\tag{6.6}
\]
Denote
\[
A = h_t - Lh - G_1 (x, t) \nabla h - G_2 (x, t) h - G_3 (x, t, \nabla h, h),
B = v_{1t} - Lv_1 - G (\nabla v_1, v_1, x, t).
\]
Then (6.6) implies that
\[
v_{2t} - Lv_2 - G (\nabla v_2, v_2, x, t) = A + B.
\]
Hence,
\[
(v_{2t} - Lv_2 - G (\nabla v_2, v_2, x, t))^2 - (v_{1t} - Lv_1 - G (\nabla v_1, v_1, x, t))^2
\]
\[
= (A + B)^2 - B^2 = 2BA + A^2
\]
\[
= 2 [v_{1t} - Lv_1 - G (\nabla v_1, v_1, x, t)] [(h_t - Lh) - G_1 (x, t) \nabla h - G_2 (x, t) h]
\tag{6.7}
\]
\[
-2G_3 (x, t, \nabla h, h) [v_{1t} - Lv_1 - G (\nabla v_1, v_1, x, t)]
\]
\[
+ [(h_t - Lh) - G_1 (x, t) \nabla h - G_2 (x, t) h - G_3 (x, t, \nabla h, h)]^2
\]
\[
= \text{Lin} (h) (x, t) + \text{Nonlin} (h) (x, t),
\]
where Lin $(h)$ and Lin $(h)$ denote respectively linear and nonlinear expressions of (6.7) with respect to $h$. The linear term is:
\[
\text{Lin} (h) (x, t)
\tag{6.8}
\]
\[
= 2 [v_{1t} - Lv_1 - G (\nabla v_1, v_1, x, t)] [(h_t - Lh) - G_1 (x, t) \nabla h - G_2 (x, t) h].
\]
We represent the term Nonlin $(h) (x, t)$ in (6.7) as
\[
\text{Nonlin} (h) = [(h_t - Lh) - G_1 (x, t) \nabla h - G_2 (x, t) h - G_3 (x, t, \nabla h, h)]^2
\]
\[
-2G_3 (x, t, \nabla h, h) [v_{1t} - Lv_1 - G (\nabla v_1, v_1, x, t)]
\]
\[ \begin{aligned}
&= (h_t - Lh)^2 - 2 (h_t - Lh) [G_1 (x, t) \nabla h + G_2 (x, t) h + G_3 (x, t, \nabla h, h)] \\
&\quad + [G_1 (x, t) \nabla h + G_2 (x, t) h + G_3 (x, t, \nabla h, h)]^2 \\
&\quad - 2G_3 (x, t, \nabla h, h) [\nu_1 - L\nu_1 - G (\nabla \nu_1, \nu_1, x, t)] .
\end{aligned} \]

Hence, using Cauchy--Schwarz inequality, (6.4) and (6.5), we obtain

\[ \text{Nonlin} (h) (x, t) \geq \frac{1}{2} (h_t - Lh)^2 - C_2 \left( (\nabla h)^2 + h^2 \right) , \forall (x, t) \in Q_T. \quad (6.9) \]

Also, by (5.6) and (6.4)--(6.7)

\[ |\text{Nonlin} (h)| (x, t) \leq C_2 \|h\|^2_{C^0(\Omega_T)} \leq C_2 \|h\|^2_{H^2(Q_T)} , \forall (x, t) \in Q_T. \quad (6.10) \]

Thus, (5.10), (6.7) and (6.8) imply that

\[ I_{\alpha, \nu} (\nu_1 + h) - I_{\alpha, \nu} (\nu_1) \\
= \exp (-2 (\tau + 1)\nu) \int_{Q_T} \text{Lin} (h) \exp (2 (t + 1)\nu) \, dx \, dt + 2\alpha (\nu_1, h)_k \\
+ \exp (-2 (\tau + 1)\nu) \int_{Q_T} \text{Nonlin} (h) \exp (2 (t + 1)\nu) \, dx \, dt + \alpha \|h\|^2_{H^2(Q_T)} , \]

where \((., .)_k\) is the scalar product in \(H^2 (Q_T)\). Assuming temporary that \(h\) is an arbitrary function of \(H^2 (Q_T)\), consider the expression \(X (h)\),

\[ X (h) = \exp (-2 (\tau + 1)\nu) \int_{Q_T} \text{Lin} (h) \exp (2 (t + 1)\nu) \, dx \, dt + 2\alpha (\nu_1, h)_k . \quad (6.12) \]

It follows from (5.9), (6.5) and (6.8) that \(X : H^2 (Q_T) \to \mathbb{R}\) is a bounded linear functional. Hence, by Riesz theorem there exists a unique function \(\check{X} \in H^2 (Q_T)\) such that

\[ X (h) = \left( \check{X}, h \right)_k . \quad (6.13) \]

At the same time, it follows from (6.10)--(6.13) that

\[ I_{\alpha, \nu} (\nu_1 + h) - I_{\alpha, \nu} (\nu_1) = \left( \check{X}, h \right)_k + o \left( \|h\|_{H^2(Q_T)} \right) , \quad (6.14) \]

\[ \lim_{\|h\|_{H^2(Q_T)} \to 0} \left( o \left( \|h\|_{H^2(Q_T)} \right) \right) = 0. \]

Hence, \(\check{X}\) is the Fréchet derivative \(I'_{\alpha, \nu} (\nu_1) \in H^2 (Q_T)\) of the functional \(I_{\alpha, \nu} (\nu)\) at the point \(\nu_1\),

\[ \left( \check{X}, h \right)_k = I'_{\alpha, \nu} (\nu_1) (h) , \forall h \in H^2 (Q_T) . \quad (6.15) \]

We now come back again to the case when \(h \in B (2\mathcal{R})\) as in (6.1). Using (6.9) and (6.11)--(6.15), we obtain
\[ I_{\alpha,\nu} (v_1 + h) - I_{\alpha,\nu} (v_1) - I'_{\alpha,\nu} (v_1) \langle h \rangle \]
\[
\geq \frac{1}{2} \exp \left( -2 (\tau + 1)^\nu \right) \int_{Q_T} \left( h_t - Lh \right)^2 \exp \left( 2 (t + 1)^\nu \right) \, dx \, dt \tag{6.16}
\]
\[- C_2 \exp \left( -2 (\tau + 1)^\nu \right) \int_{Q_T} \left( (\nabla h)^2 + h^2 \right) \exp \left( 2 (t + 1)^\nu \right) \, dx \, dt + \alpha \| h \|_{H^\nu (Q_T)}^2. \]

We now use the Carleman estimate (2.2). Recalling that \( h(x, T) = 0 \) and using (6.16), we obtain

\[ I_{\alpha,\nu} (v_1 + h) - I_{\alpha,\nu} (v_1) - I'_{\alpha,\nu} (v_1) \langle h \rangle \]
\[
\geq \frac{\mu_1}{12} \sqrt{\nu} \exp \left( -2 (\tau + 1)^\nu \right) \int_{Q_T} (\nabla h)^2 \exp \left( 2 (t + 1)^\nu \right) \, dx \, dt \tag{6.17}
\]
\[ + \frac{\nu^2}{24} \exp \left( -2 (\tau + 1)^\nu \right) \int_{Q_T} h^2 \exp \left( 2 (t + 1)^\nu \right) \, dx \, dt \]
\[- C_2 \exp \left( -2 (\tau + 1)^\nu \right) \int_{Q_T} \left( (\nabla h)^2 + h^2 \right) \exp \left( 2 (t + 1)^\nu \right) \, dx \, dt \]
\[- \mu_2 \nu^2 \exp \left( -2 (\tau + 1)^\nu \right) \| \nabla h (x, 0) \|_{L^2 (\Omega)}^2 + \alpha \| h \|_{H^\nu (Q_T)}^2. \]

By (5.6)
\[
\| \nabla h (x, 0) \|_{L^2 (\Omega)}^2 \leq C_2 \| h \|_{H^\nu (Q_T)}^2. \tag{6.18}
\]

Choose the number \( \nu_3 \geq \nu_0 > 1 \) depending on the same parameters as those listed in (5.13) such that
\[
\mu_1 \sqrt{\nu_3} \geq 24 C_2 \text{ and } \nu_3^2 \geq 48 C_2. \tag{6.19}
\]

Recall that \( C_2 > 0 \) denotes different constants depending on parameters listed in (5.13). Hence, (5.11), (5.12), (5.14) and (6.16)–(6.19) imply that for all \( \nu \geq \nu_3 \)
\[ I_{\alpha,\nu} (v_1 + h) - I_{\alpha,\nu} (v_1) - I'_{\alpha,\nu} (v_1) \langle h \rangle \]
\[
\geq \frac{\mu_1}{24} \sqrt{\nu} \exp \left( -2 (\tau + 1)^\nu \right) \int_{Q_T} (\nabla h)^2 \exp \left( 2 (t + 1)^\nu \right) \, dx \, dt \tag{6.16}
\]
\[ + \frac{\nu^2}{48} \exp \left( -2 (\tau + 1)^\nu \right) \int_{Q_T} h^2 \exp \left( 2 (t + 1)^\nu \right) \, dx \, dt \]
\[ + \| h \|_{H^\nu (Q_T)}^2 \left( \alpha - C_2 \mu_2 \nu^2 \exp \left( -2 (\tau + 1)^\nu \right) \right) \]
\[
\geq \frac{\mu_1}{24} \sqrt{\nu} \exp \left( -2 (\tau + 1)^\nu \right) \int_{Q_T} (\nabla h)^2 \exp \left( 2 (t + 1)^\nu \right) \, dx \, dt
\]

27
\[ + \frac{\nu^2}{48} \exp\left(-2 (\tau + 1)^\nu\right) \int_{Q_T} h^2 \exp\left(2 (t + 1)^\nu\right) \, dx \, dt + \frac{\alpha}{2} \|h\|^2_{H^2(Q_T)} \]
\[ \geq C_2 \|h\|^2_{H^2(Q_T)} + \frac{\alpha}{2} \|h\|^2_{H^2(Q_T)}, \]
which implies (5.15). \[ \square \]

6.2. Proof of theorem 5.4

Denote
\[ I_{\alpha,\nu}^0(v) = \exp\left(-2 (\tau + 1)^\nu\right) \int_{Q_T} (v_t - L v - G (\nabla v, v, x, t))^2 \exp\left(2 (t + 1)^\nu\right) \, dx \, dt. \] (6.20)

By (5.10) and (6.20)
\[ I_{\alpha,\nu}(v) = I_{\alpha,\nu}^0(v) + \alpha \|v\|^2_{H^2(Q_T)}. \] (6.21)

By (5.5)
\[ G (\nabla v, v, x, t) = Lg + F (\nabla v + \nabla g, v + g, x, t) \]
\[ = Lg^* + (Lg - Lg^*) + F [(\nabla v + \nabla g^*) + (\nabla g - \nabla g^*) + (v + g^*) + (g - g^*), x, t]. \]

Hence, using the multidimensional analog of Taylor formula [31] and (5.20), we obtain similarly with (6.3)–(6.5)
\[ G (\nabla v, v, g, x, t) = Lg^* + F (\nabla v^* + \nabla g^*, v^* + g^*, x, t) + P(x, t) \] (6.22)
\[ = G (\nabla v^*, v^*, g^*, x, t) + P(x, t), \]
where the function \( P(x, t) \) is such that
\[ \|P\|_{L^1(Q_T)} \leq C_2 \delta. \] (6.23)

Since \( v_t - L v^* - G (\nabla v^*, v^*, g^*, x, t) = 0 \), then, using (6.23), we obtain
\[ I_{\alpha,\nu}^0(v^*) = \exp\left(-2 (\tau + 1)^\nu\right) \]
\[ \times \int_{Q_T} [v_t^* - L v^* - G (\nabla v^*, v^*, g^*, x, t) + P(x, t)]^2 \exp\left(2 (t + 1)^\nu\right) \, dx \, dt \]
\[ = \exp\left(-2 (\tau + 1)^\nu\right) \int_{Q_T} P^2(x, t) \exp\left(2 (t + 1)^\nu\right) \, dx \, dt \]
\[ \leq C_2 \exp\left(2 (T + 1)^\nu\right) \delta^2. \]

Hence, (6.21) implies that
\[ I_{\alpha,\nu}(v^*) \leq C_2 \left( \exp\left(2 (T + 1)^\nu\right) \delta^2 + \alpha \right). \] (6.24)

Next, by (5.15)
\[ I_{\alpha,\nu}(v^*) - I_{\alpha,\nu}(v_{\min}) \leq I_{\alpha,\nu}^0(v_{\min}) (v^* - v_{\min}) \]
\begin{align}
&\geq C_2 \left\| \mathbf{v}^* - v_{\min} \right\|_{H^1(\Omega_T)}^2 + \frac{\alpha}{2} \left\| \mathbf{v}^* - v_{\min} \right\|_{H^2(\Omega_T)}^2. 
\end{align}

By (5.17) \(-I_{\Delta, \nu} (v_{\min}) (\mathbf{v}^* - v_{\min}) \leq 0\). Hence, (6.24) and (6.25) imply that
\begin{align}
\left\| \mathbf{v}^* - v_{\min} \right\|_{H^1(\Omega_T)}^2 \leq C_2 \left( \exp \left( 2 (T + 1)^\nu \right) \delta^2 + \alpha \right). 
\end{align}

Recall that the numbers \( \nu = \nu(\delta) \) and \( \alpha = \alpha(\delta) \) are defined in (4.33) and (5.22) respectively. These choices of \( \nu(\delta) \) and \( \alpha(\delta) \) are possible since (5.21) holds and \( \delta \in (0, \delta_0) \). Thus, condition (5.14) of theorem 5.1 imposed on \( \alpha \) is in place. Hence, (4.32) and (4.34) hold. Hence, using (3.2), (4.32)–(4.34), and (6.26), we obtain
\begin{align}
\left\| \mathbf{v}^* - v_{\min} \right\|_{H^1(\Omega_T)}^2 \leq C_2 \exp \left[ \frac{2}{\bar{T}} \left( \ln \left( \frac{\delta}{\delta_0} \right)^{-1} \right)^2 \right],
\end{align}
which implies (5.23). To prove (5.24), we use the triangle inequality,
\begin{align}
&\left\| \mathbf{v}^* - v_n \right\|_{H^1(\Omega_T)}^2 \leq \left\| \mathbf{v}^* - v_{\min} \right\|_{H^1(\Omega_T)}^2 + \left\| v_n - v_{\min} \right\|_{H^1(\Omega_T)}^2
\end{align}
Using (5.19), (5.23) and (6.27), we obtain (5.24).

Acknowledgments

The work of M V Klibanov was supported by US Army Research Laboratory and US Army Research Office grant W911NF-19-1-0044. The work of A G Yagola was supported by Russian Foundation of Basic Research grant 17-01-00159.

ORCID iDs

Michael V Klibanov \( \text{https://orcid.org/0000-0001-5179-6879} \)

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 2017 Convergent algorithm based on carleman estimates for the recovery of a potential in the wave equation *SIAM J. Numer. Anal.* **55** 1578–613
[3] Beilina L and Klibanov M V 2012 *Approximate Global Convergence and Adaptivity for Coefficient Inverse Problems* (New York: Springer)
[4] Beilina L and Klibanov M V 2015 Globally strongly convex cost functional for a coefficient inverse problem *Nonlinear Anal.: Real World Appl.* **22** 272–8
[5] 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
[6] Bourgeois L, Pomonarev D and Dardé J 2019 An inverse obstacle problem for the wave equation in a finite time domain *Inverse Problems Imaging* **13** 377–400
[7] Clason C and Klibanov M V 2007 The quasi-reversibility method for thermoacoustic tomography in a heterogeneous medium *SIAM J. Sci. Comput.* **30** 1–23
[8] Friedman A 1964 *Partial Differential Equations of Parabolic Type* (Englewood Cliffs, NJ: Prentice-Hall)
[9] Hao D N, Duc N V and Thang N V 2018 Backward semi-linear parabolic equations with time-dependent coefficients AMD local Lipschitz source Inverse Problems 34 05510
[10] Isakov V 2006 Inverse Problems for Partial Differential Equations 2nd edn (New York: Springer)
[11] Kaltenbacher B and Rundell W 2019 Regularization of a backwards parabolic equation by fractional operators Inverse Problems Imaging 13 401–30
[12] Klibanov M V and Santos F 1991 A computational quasi-reversibility method for Cauchy problems for Laplace’s equation SIAM J. Appl. Math. 51 1653–75
[13] Klibanov M V and Ioussoupova O V 1995 Uniform strict convexity of a cost functional for three-dimensional inverse scattering problem SIAM J. Math. Anal. 26 147–79
[14] Klibanov M V 1997 Global convexity in a three-dimensional inverse acoustic problem SIAM J. Math. Anal. 28 1371–88
[15] Klibanov M V and Timonov A 2004 Carleman Estimates for Coefficient Inverse Problems and Numerical Applications (Utrecht: VSP)
[16] Klibanov M V 2015 Carleman estimates for the regularization of ill-posed Cauchy problems Appl. Numer. Math. 94 46–74
[17] Klibanov M V 2015 Carleman weight functions for solving ill-posed Cauchy problems for quasilinear PDEs Inverse Problems 31 125007
[18] Klibanov M V and Kamburg V G 2016 Globally strictly convex cost functional for an inverse parabolic problem Math. Methods Appl. Sci. 39 930–40
[19] Klibanov M V, Kuzhuget A V and Golubnichiy K V 2016 An ill-posed problem for the Black–Scholes equation for a profitable forecast of prices of stock options on real market data Inverse Problems 32 015010
[20] Klibanov M V, Koshev N A, Li J and Yagola A G 2016 Numerical solution of an ill-posed Cauchy problem for a quasilinear parabolic equation J. Inverse Ill-Posed Problems 24 761–76
[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 and Nguyen D-L 2019 Convexification method for an inverse scattering problem and its performance for experimental backscatter data for buried targets SIAM J. Imaging Sci. 12 576–603
[23] Klibanov M V, Li J, Li and Zhang W 2019 Convexification of electrical impedance tomography with restricted Dirichlet-to-Neumann map data Inverse Problems 35 035005
[24] Klibanov M V, Li J and Zhang W 2018 Convexification for the inversion of a time dependent wave front in a heterogeneous medium (arXiv: 1812.11281)
[25] Klibanov M V and Nguyen L H 2019 PDE-based numerical method for a limited angle x-ray tomography Inverse Problems 35 045009
[26] Lavrentiev M M, Romanov V G and Shishatskii S P 1986 Ill-Posed Problems of Mathematical Physics and Analysis (Providence, RI: American Mathematical Society)
[27] Lattes R and Lions J L 1969 The Method of Quasireversibility: Applications to Partial Differential Equations (New York: Elsevier)
[28] Scales J A, Smith M L and Fischer T L 1992 Global optimization methods for multimodal inverse problems J. Comput. Phys. 103 258–68
[29] Tikhonov A N, Goncharsky A V, Stepanov V V and Yagola A G 1995 Numerical Methods for the Solution of Ill-Posed Problems (London: Kluwer)
[30] Tuan N H, Khoa V A and Au V V 2019 Analysis of a quasi-reversibility method for a terminal value quasilinear parabolic problem with measurements SIAM J. Math. Anal. 51 60–85
[31] Vainberg M M 1973 Variational Method and Method of Monotone Operators in the Theory of Nonlinear Equations (Washington, DC: Wiley)